

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
15 July 2004 (15.07.2004)

PCT

(10) International Publication Number  
WO 2004/058055 A2

(51) International Patent Classification: A61B

Panoramic Lane, San Diego, CA 92121 (US). DAHLEN, Jeffrey, R. [US/US]; 10555 Kemmerion Road, San Diego, CA 92126 (US).

(21) International Application Number: PCT/US2003/041426

(74) Agents: WARBURG, Richard, J. et al.; FOLEY & LARDNER, P.O. Box 80278, San Diego, CA 92138-0278 (US).

(22) International Filing Date: 23 December 2003 (23.12.2003)

(25) Filing Language: English

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SG, SI, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(26) Publication Language: English

(30) Priority Data:  
60436,392 24 December 2002 (24.12.2002) US  
10/331,127 27 December 2002 (27.12.2002) US  
10/410,572 8 April 2003 (08.04.2003) US

(71) Applicant (for all designated States except US): BIOSITE INCORPORATED [US/US]; 11030 Roselle Street, San Diego, CA 92121 (US).

(84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

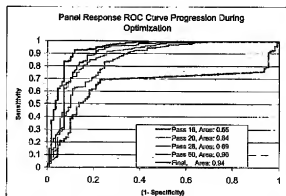
Published:

(75) Inventors/Applicants (for US only): ANDERBERG, Joseph, Michael [US/US]; 470 Delage Ct., Encinitas, CA 92024 (US). BUECHLER, Kenneth, F. [US/US]; P.O. Box 77, Rancho Santa Fe, CA 92067 (US). MCPHERSON, Paul, H. [US/US]; 1449 Elva Court, Encinitas, CA 92024 (US). KIRCHICK, Howard, J. [US/US]; 5449

— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: METHOD AND SYSTEM FOR DISEASE DETECTION USING MARKER COMBINATIONS



(57) Abstract: The present invention relates to methods and system for the diagnosis diseases or conditions. In a particular aspect, a disclosed method for determining a panel includes calculating a panel response for each patient in a set of diseased patients and in a set of non-diseased patients. The panel response is a function of the value of each of a plurality of markers in a panel of markers. The method also includes calculating a value for an objective function. The objective function is indicative of the effectiveness of the panel. The steps of calculating a panel response for each patient and calculating a value for an objective function are iterated by varying at least one of the parameters relating to the panel response function and a sense of each marker to facilitate optimization of the objective function. The objective function may be a measure of an overlap of panel responses of diseased patients and panel responses of non-diseased patients. The contribution of each marker to the objective function may be determined, and the panel size may be reduced by removing the poorest markers. Thus, an optimum panel of markers and an optimal panel response function for the diagnosis of a disease or condition may be determined.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

## METHOD AND SYSTEM FOR DISEASE DETECTION USING MARKER COMBINATIONS

### FIELD OF THE INVENTION

[0001] The present invention relates to the identification and use of diagnostic markers for various diseases or conditions. More particularly, the invention relates to methods and systems for identifying and utilizing panel of markers for detection of one or more particular diseases or conditions.

### BACKGROUND OF THE INVENTION

[0002] The background of the invention is provided to aid the reader in understanding the invention and is not admitted to describe or constitute prior art to the present invention.

[0003] The clinical presentation of certain diseases can often be strikingly similar, even though the underlying diseases, and the appropriate treatments to be given to one suffering from the various diseases, can be completely distinct. For example, subjects may present in an urgent care facility exhibiting a deceptively simple constellation of apparent symptoms (*e.g.*, fever, shortness of breath, dizziness, headache) that may be characteristic of a variety of unrelated conditions. Diagnostic methods often involve the comparison of symptoms and/or diagnostic test results known to be associated with one or more diseases that exhibit a similar clinical presentation to the symptoms and/or diagnostic results exhibited by the subject, in order to identify the underlying disease or condition present in the subject.

[0004] The acuteness or severity of the symptoms often dictates how rapidly a diagnosis must be established and treatment initiated. For example, immediate diagnosis and care of a patient experiencing a variety of acute conditions can be critical. *See, e.g.*, Harris, *Aust. Fam. Physician* 31: 802-06 (2002) (asthma); Goldhaber, *Eur. Respir. J. Suppl.* 35: 22s-27s (2002) (pulmonary embolism); Lundergan *et al.*, *Am. Heart J.* 144: 456-62 (2002) (myocardial infarction). However, even in cases where the apparent symptoms appear relatively stable, rapid diagnosis, and the rapid initiation of treatment, can provide both relief from immediate discomfort and advantageous improvement in prognosis.

[0005] Recently, workers seeking to provide rapid diagnostic methods for various diseases or conditions have sought to identify "markers" for diseases; that is, molecules that are present in a sample obtained from a subject suffering from a disease of interest in an amount that differs from the amount present in a sample from a "normal," non-diseased subject.

[0006] Diagnoses of many diseases or conditions, such as cardiovascular disease and stroke, for example, are performed by measurement of the levels of particular markers in a patient. Often, however, a single marker is generally incapable of providing clinical utility because its value does not provide a means of confidently distinguishing between a diseased patient and a non-diseased patient.

[0007] As an example, Figure 1 illustrates that the levels of a particular marker expressed in a diseased and a non-diseased population. As shown in the figure, the marker levels in these two populations may be distributed over broad ranges in a distribution pattern. Although the diseased population in this example generally may exhibit higher or lower levels for the marker than the non-diseased population, substantial portions of each population fall within a region of overlapping values. Thus, definitive or confident diagnosis of a disease or a condition based on the measurement of this single marker may be impossible. Traditionally one chooses a cutoff value in the overlap region. The cutoff is chosen to optimize the number of false positive versus the number of false negatives. In practice physicians often treat a patient based on where they fall relative to the cutoff. They often do not consider how close the patient is to the cutoff.

[0008] The effectiveness of a test having such an overlap is often expressed using a ROC (Receiver Operating Characteristic) curve. Other measures, such as positive predictive value (PPV) and negative predictive value (NPV) may also be used as a measure of the effectiveness of the test. ROC curves are well known to those skilled in the art. Thus, the details pertaining to ROC curves are beyond the scope of this document, however there is a brief description below. Further, reference may be made to Zweig, M.H. & Campbell, C.C., Clin Chem 39, 561-577 (1993) and Henderson, A.R., Ann. Clin. Biochem 30, 521-539 (1993).

[0009] Figure 3 illustrates an example of a ROC curve for the marker level distributions of Figure 1. The ROC curve shows the trade off between the sensitivity and

specificity of a marker. The sensitivity is a measure of the ability of the marker to detect the disease, and the specificity is a measure of the ability of the marker to detect the absence of the disease. The horizontal axis of the ROC curve represents (1- specificity), which increases with the rate of false positives. The vertical axis of the curve represents sensitivity, which increases with the rate of true positives. Thus, for a particular cutoff selected, the values of specificity and sensitivity may be determined. The right hand end of the curve is the minimum cutoff, the left hand end of the curve is the maximum cutoff. As the cutoff is changed to increase specificity, sensitivity usually is reduced and vice versa. The area under the ROC curve is a measure of the utility of the measured marker level in the correct identification of one or more diseases or conditions. Thus, the area under the ROC curve can be used to determine the effectiveness of the test. Note the area is independent of the cutoff value.

[0010] Panels of multiple markers may improve the likelihood of an accurate diagnosis. The multiple marker "panel" for a particular disease is preferably selected such that a particular "profile" of marker levels is specific for that disease and capable of clearly distinguishing disease from non-disease. However, methods for identifying such panels, and the particular "profiles" that provide clinical utility, are typically empirical in nature, relying on trial-and-error. Furthermore, because the computational complexity involved in identifying suitable diagnostic thresholds and/or profiles increases as the number of markers in a potential panel increase, marker panels typically involve only a few markers. Searching for an effective panel from among a large number of markers can become the computational equivalent of finding a needle in a haystack. For example, often one might look for elevation of 4 of 6 markers, or more generally  $n$  of  $m$  markers, to define a positive state. In this example the cutoff values for each marker are chosen, then the data analyzed to see how effective the test is. This is repeated for different number of elevated markers, cutoffs and markers. In this example, all markers are treated with equal importance, there is no method to adjust the relative importance.

#### BRIEF SUMMARY OF THE INVENTION

[0011] The method disclosed in this document provides a means to systematically find the optimal markers and panels of markers to distinguish (compare) non-disease from disease, and it also optimizes the way in which the marker values are used. A first step to

simplify the problem of defining a marker or a panel of markers is defining an 'objective function'. An objective function is a scalar function, and will represent the effectiveness of the test for diagnosis of non-disease from disease. For example, rather than requiring  $n$  elevated markers to define a positive state and then quantifying the effectiveness of this algorithm, one can generate a ROC curve from the number of elevated markers, and use the area under the ROC curve ("the ROC curve area") to define the effectiveness of the test. By using the ROC curve area as the effectiveness of the test, the optimization problem has been simplified. This is because the search space has been reduced since there is no need to calculate the effectiveness associated with each of the  $m$  values for  $n$  elevated markers. In this example, the number of elevated markers can be thought of as a concentration for the ROC curve, but as described above, the selection of the cutoff concentration is not required to determine if a test will be effective. Another step to simplify the problem of defining a marker or a panel of markers may be to define a systematic way to find the best way to use the markers. Without this it is very difficult to find the best markers because one needs to distinguish the markers and how to use them. A systematic method to find the best way to use the markers is to combine all the values into one result, the "panel response". Functional forms of the panel response can be selected in a variety of manners described herein. Once this is done search routines can be employed to find the panel response function to maximize or minimize the objective function for a set of markers.

[0012] The method may also include a technique for determining the relative importance of the markers in the set, and subsequently determine the optimum markers to use, for example, in a panel of  $n$  markers.

[0013] In addition to measured marker levels, other information including a patient's history, sex, age, race, and other factors may also require consideration. In this regard, embodiments of the disclosed method may accommodate such factors as markers.

[0014] Specifically, certain disclosed embodiments of the present invention relate to the identification and use of diagnostic markers for cardiac diseases and stroke and cerebral injury. Generally, the methods and systems described herein can meet the need in the art for the development of an effective panel of markers for the accurate diagnosis of a selected disease or condition. More generally, the disclosed methods and systems may be

used to develop criteria for distinguishing members of two or more groups for whom the distribution of certain characteristics are known.

[0015] In a first aspect, the invention discloses a method of identifying a panel of markers for determining a diagnosis of a disease or a condition, or of a prognosis (that is a risk of some future outcome). The method includes calculating a panel response for each patient in two sets of patients, referred to for convenience as "diseased" patients and "non-diseased" patients. The panel response is a function of value of each of a plurality of markers in a panel of markers.

[0016] The term "panel" as used herein refers to a set of markers. The panel may include any practical number of markers appropriate for use with the diagnosis of the particular one or more diseases or conditions.

[0017] The term "marker" as used herein refers to proteins, polypeptides, nucleic acids, bacteria, viruses, prions, small molecules and the like, to be used as targets for screening test samples obtained from subjects. "Proteins, polypeptides, or small molecules" used as markers in the present invention are contemplated to include any fragments thereof, in particular, immunologically detectable fragments. "Marker", as used herein, may include derived markers as defined below, and may also include such characteristics as patient's history, age, sex and race, for example. Certain markers are also known in the field as "analytes". A marker is said to be a specific marker of the disease if only the presence or absence of the target disease condition influences its value. A marker is said to be a nonspecific marker of the disease if many disease conditions influence its value. An example of a specific marker is cardiac TnI, which, when elevated above about 1 ng/ml is specific to myocardial infarction. An example of a non specific marker is C-reactive protein ("CRP"), which is elevated in conditions that promote the inflammatory response.

[0018] The term "related marker" as used herein refers to one or more fragments of a particular marker or its biosynthetic parent that may be detected as a surrogate for the marker itself or as independent markers. For example, human BNP is derived by proteolysis of a 108 amino acid precursor molecule, referred to hereinafter as BNP<sub>1-108</sub>. Mature BNP, or "the BNP natriuretic peptide," or "BNP-32" is a 32 amino acid molecule representing amino acids 77-108 of this precursor, which may be referred to as BNP<sub>77-108</sub>.

The remaining residues 1-76 are referred to hereinafter as BNP<sub>1-76</sub>. Additionally, related markers may be the result of covalent modification of the parent marker, for example by oxidation of methionine residues, ubiquitination, cysteinylolation, nitrosylation, glycosylation, etc.

[0019] Preferably, the methods described hereinafter utilize one or more markers that are derived from the subject. The term "subject-derived marker" as used herein refers to protein, polypeptide, phospholipid, nucleic acid, prion, or small molecule markers that are expressed or produced by one or more cells of the subject. The presence, absence, amount, or change in amount of one or more markers may indicate that a particular disease is present, or may indicate that a particular disease is absent. Additional markers may be used that are derived not from the subject, such as molecules expressed by pathogenic or infectious organisms that are correlated with a particular disease, race, time since onset, sex, etc. Such markers are preferably protein, polypeptide, phospholipid, nucleic acid, prion, or small molecule markers that identify the infectious diseases described above.

[0020] The term "diagnosis" as used herein refers to methods by which the skilled artisan can estimate and/or determine whether or not a patient is suffering from a given disease or condition. The skilled artisan often makes a diagnosis on the basis of one or more diagnostic markers, the presence, absence, or amount of which may be indicative of the presence, severity, or absence of the condition. In addition to markers, other tests, such as ECG, Echo, and MRI, and other factors, such as patient's history, sex, age, and race, may also be used in making the diagnosis. As used herein, the term "markers" also includes these other tests and other factors. While the instant specification describes the invention in terms of diagnosis of disease, the methods described herein are equally applicable to identifying markers for use in prognosis. A "prognosis" may be determined by measuring one or more markers, the presence or amount of which in a subject (or a sample obtained from the subject) signal a probability that a given course or outcome will occur.

[0021] The term "panel response" as used herein refers to a scalar function or its value, which is a function of the marker values of the panel. Most generally, the panel response is a function of the marker values ( $M_{1-n}$ ), written as  $PR = f(M_{1-n})$ . In a preferred embodiment the panel response is a summation over indicator values (I) of each



marker. The indicator value is generally a function of the marker value. This can be represented as  $PR = \sum_{\text{Marker } i} I_i(M_i) \cdot W_i$ , where  $I_i$  is a function of the marker value  $M_i$ ,  $W_i$  is a weighting coefficient that scales the indicator function. For definitive purposes, in this document it will be assumed that the panel response is scaled such that all values are between 0 and 1, but other increments can apply.

[0022] The terms "diseased" and "non-diseased" as used herein refer to two populations that differ in terms of a disease characteristic of interest. For example, a "diseased" population may be a population suffering from a stroke, while a "non-diseased" population may be a population not suffering from a stroke. These are simply labels, and the skilled artisan will understand that one could just as easily label the population suffering from a stroke "non-diseased" and the other population "diseased" without varying from the teachings set forth herein. In the case of a prognosis, the two populations may each be suffering from the same disease, but separated into populations on the basis of one group having a particular outcome (e.g., death) and the second group not having that outcome. The set of diseased patients and set of non-diseased patients may include patients whose state, whether diseased or non-diseased, has been confirmed and for whom marker levels are available for one or more markers.

[0023] The term "marker value" as used herein refers to a numeric value, such as a value representing the result of an assay of the marker. For example, the marker value may be expressed in units of concentration or number. When the marker represents characteristics such as a patient's history, then the value may be a numeric representation, or mapping, of the history information.

[0024] The term "derived marker" as used herein refers to a value that is a function of one or more measured markers. For example, derived markers may be related to the change over a time interval in one or more measured marker values, may be related to a ratio of measured marker values, may be a marker value at a different measurement time, or may be a complex function such as a panel response function.

[0025] The method further comprises calculating a value for an objective function, the objective function being indicative of an effectiveness of the panel.

[0026] The term "objective function" as used herein refers to a scalar function or its value, which may be a function of the plurality of panel responses and known disease states or diagnoses of a collection of patient samples. The objective function is a measure of the clinical effectiveness of the test, or the ability to distinguish disease from non-disease. An example of an objective function is the area under the ROC curve. The objective function may be related to the amount of overlap between the diseased and non-diseased panel response values. The objective function is a scalar value, which is indicative of the effectiveness of the panel. The objective function may be defined by a user as a function of various outputs, such as ROC curve features defined below, of the panel responses for the groups of patients.

[0027] The method of the first aspect of the invention also comprises iterating the calculating a panel response for each patient and calculating a value for an objective function by varying at least one of parameters relating to the panel response function and a sense of each marker to facilitate optimization of the objective function.

[0028] "Iterating" may include repeating the steps with variations in the inputs, where the variations may be dependant on the outputs of the previous iteration. "Varying" may include tweaking a parameter by either a predetermined amount, an amount dependant on an output of the previous iteration or a random amount.

[0029] The term "sense" as used herein refers to the direction of the response of a marker with disease state. If a marker value is elevated in diseased patients relative to non-diseased patients, then the marker is said to have a positive sense. If the marker value is lower in diseased patients relative to non-diseased patients then the marker is said to have a negative sense. If the probability of a finding the marker value near some specific value is elevated in diseased patients relative to non-diseased patients, the sense is said to be positive. If the probability of a finding a marker value near some specific value is reduced in diseased patients relative to non-diseased patients, the sense is said to be negative. One skilled in the art will recognize that it is trivial to invert functions or map the marker value such that a negative sense marker can be analyzed in the same way as a positive sense marker. Moreover, as discussed above, the labeling of a population as "diseased" or "non-diseased" is merely a question of labeling of two populations for the presence or absence of a characteristic of interest. Throughout this document the marker sense is described as

positive. This is for conciseness only, all concepts and claims can apply to both negative and positive sense markers, and both positive and negative senses are implicitly included.

[0030] The term "parameters" as used herein refers to coefficients, powers, etc. of a function that may be varied to modify the functional value. For example, if the function is a ramp function, the low threshold and the high threshold, may be two parameters that are varied. If the function is a Gaussian the width and location may be two parameters that are varied. The optimization process will modify one or more of the parameters of the panel response function, which in one embodiment may include all of the parameters of the used indicator functions and weighting coefficients.

[0031] According to another aspect of the invention, a system for identifying a panel of markers for diagnosis/prognosis of a disease or a condition includes means for calculating a panel response for each patient in a set of diseased patients and in a set of non-diseased patients. In one embodiment the panel response is a function of a value of each of a plurality of markers in a panel of markers. The means for calculating may be a central processing unit (CPU), as may be available on a desktop computer, a laptop computer, a workstation or a mainframe, for example.

[0032] The system further includes means for calculating a value for an objective function. The objective function is indicative of the effectiveness of the panel. In certain embodiments, an objective function may be a measure of overlap of panel responses of diseased patients and panel responses of non-diseased patients.

[0033] Further, the system includes means for iteratively activating the means for calculating a panel response and the means for calculating a value for an objective function, by varying at least one of the following parameters to facilitate optimization of said objective function: parameters relating to the panel response function and a sense of each marker.

[0034] In another aspect of the invention, a program product includes machine readable program code for causing a machine to perform certain method steps. The method steps include calculating a panel response for each patient in a set of diseased patients and in a set of non-diseased patients. The panel response is a function of value of each of a plurality of markers in a panel of markers.

[0035] The method steps further include calculating a value for an objective function. The objective function is indicative of the effectiveness of the panel. Further, the method steps include iterating the steps of calculating a panel response for each patient and calculating a value for an objective function by varying at least one of the following parameters to facilitate optimization of said objective function: parameters relating to the panel response function and a sense of each marker.

[0036] In a preferred embodiment, the program product includes machine readable code embedded in a portable meter. The term "portable meter," as used herein, may include any number of devices having the ability to execute coded instructions. In a further preferred embodiment, the portable meter is a fluorometer. In an alternate embodiment, the portable meter is a reflectometer.

[0037] In a preferred embodiment, the program product includes machine readable code embedded in a computer. In a further preferred embodiment, the computer is a portable computer. In another preferred embodiment, the computer is adapted to be accessed through a network, such as a public network like the Internet.

[0038] In another preferred embodiment, the computer is adapted to be coupled to an analyzer. In a further preferred embodiment, the analyzer is an immunoassay analyzer. In an alternate embodiment, the analyzer is a single nucleotide polymorphism detector. In another embodiment, the analyzer is adapted to sort and count similar and different particles and cells.

[0039] In a preferred embodiment, the panel response is a function of the value of an indicator for each of a plurality of markers in a panel of markers and a weighting coefficient for each marker. The indicator is a mapping, for each of the plurality of markers, of marker levels. The mapping is according to an indicator function. The iterating includes varying at least one of the weighting coefficients, parameters relating to the indicator function, and a sense of each marker to facilitate optimization of the objective function.

[0040] The term "indicator function" as used herein refers to a scalar function or its value, which is a function of a marker value. The mapping is in accordance with a indicator function. The indicator function may be any function providing a value

dependent on the marker level. The indicator function may be a mapping of marker values into values that may be more closely related to the probability of diseased state at that marker value. The indicator function may be scaled such that all values are between 0 and 1. In this document it will be assumed that the indicator function is scaled such that all values are between 0 and 1. This scaling does not influence the result of the method, however in practice it does simplify some formulations. For example, to change a positive indicator function (PIF) to work with a negative sense marker the negative indicator function (NIF) may be defined as  $NIF = 1 - PIF$ .

[0041] The term "mapping" as used herein refers to a relation between a value in one domain to a value in another domain. The mapping relation may be a one-to-one relationship or a one-to-many relationship.

[0042] The term "elevation indicator function" as used herein refers to a scalar function that has a high and monotonic rate of change between low and high threshold values, and a smaller rate of change elsewhere. Examples of this type of function include step, ramp, 'S' or sigmoid functions. One skilled in the art will recognize that there are many such functions.

[0043] The term "localization indicator function" as used herein refers to a scalar function that is peaked near some expected value, and decreases when the marker value is further away from the expected value. Examples of this type of function include triangle, square, trapezoid, or Gaussian functions. One skilled in the art will recognize that there are many such functions.

[0044] The term "contribution" as used herein refers to the relative amount that a marker contributes to the objective function. The contribution may be related to the importance of a marker.

[0045] The term "test" as used herein refers to a method performed which yields an output related to a clinical outcome. A test may comprise values of 1 or more markers. A test may also be a procedure used in the determination of a panel response. Commonly, a test is also an immunoassay.

[0046] In the method the marker values may be combined into one value, the panel response. As described above, in a preferred embodiment the panel response is represented

as  $PR = \sum_{Marker} I_i(M_i) \cdot W_i$ . Choosing different functional forms for the indicator  $I$  changes

the way a marker is used. For example, when several nonspecific markers are used, then combined elevations of the markers may indicate a diseased state. The appropriate indicator functions could be elevation indicator functions as defined above. In this example, when the marker value is below a low threshold then there is little or no change in the indicator function with marker value, and when above a high threshold then again there is little or no change in indicator value. Between these thresholds the indicator value increases or decreases with marker value. One skilled in the art will recognize there are many functions that have this property. Physically one can associate the thresholds with the lower and upper end of the overlap region as illustrated in Figure 4. A panel response may be a numerical value for each patient. The range of values of the panel response may be set to any desired range. For example, the values of the panel response may be scaled to fall between zero and one.

[0047] In another embodiment the indicator function is chosen to localize the marker value. For example if a certain pattern of marker levels is associated with a disease state then the indicator function could be a localization indicator function as defined above. These functions give a high response when the marker is near the optimal value. One skilled in the art will recognize there are many functions that have this localization property. In an example using these functions, certain disease states such as unstable angina, may be an intermediate disease. A marker such as TnI is elevated by minor necrosis due to ischemia associated with unstable angina, but is elevated still further by major necrosis associated with myocardial infarction. Other markers may be elevated with unstable angina, but not elevated with myocardial infarction. The indicator function of each analyte can be different so panels can consist of markers of both types, as needed in the example above.

[0048] In a preferred embodiment the method includes utilizing a search engine to find optimal parameters for the panel response function. The search engine is able to efficiently vary parameters of the panel response until it finds a set that results in a local maximization of the objective function. Because the objective function is a measure of the effectiveness of the test, the optimized panel response may provide an improved diagnostic value.

[0049] In a preferred embodiment the method includes calculating a contribution for each marker. In another preferred embodiment the contributions of all markers are ranked, and markers with low contribution values may be removed from the panel. The entire process can be repeated with the reduced number of markers until the desired panel size and performance are achieved.

[0050] Another embodiment of the invention measures multiple markers from a patient and combines the values into a single panel response. The panel response function could be determined by the method described above. The panel value would be compared to a cutoff value, providing an effective tool to aid in the diagnosis of disease states.

[0051] In a preferred embodiment, an objective function is a measure of overlap of panel responses of diseased patients and panel responses of non-diseased patients.

[0052] According to a preferred embodiment, the calculating of a value for an objective function includes generating a receiver operating characteristic (ROC) curve for the panel response. The ROC curve is indicative of a sensitivity of the panel response as a function of one minus a specificity of the panel response. ROC curves are well-known to those skilled in the art and are further described below.

[0053] In various aspects, multiple determination of the marker panels described herein can be made, and a temporal change in the markers can be used to rule in or out one or more diagnoses or prognoses. For example, one or more markers may be determined at an initial time, and again at a second time. In such embodiments, an increase in the marker from the initial time to the second time may be diagnostic of a particular disease, or indicate a particular prognosis. Likewise, a decrease in the marker from the initial time to the second time may be indicative of a particular disease, or of a particular prognosis.

[0054] In yet other embodiments, multiple determinations of marker panels can be made, and a temporal change in the marker can be used to monitor the efficacy appropriate therapies. In such an embodiment, one might expect to see a decrease or an increase in the marker(s) over time during the course of effective therapy.

[0055] In yet a further aspect, the invention relates to devices for analyzing the marker panels described herein. Such devices preferably contain a plurality of discrete, independently addressable locations, or "diagnostic zones," each of which is related to a

particular marker of interest. Following reaction of a sample with the devices, a signal is generated from the diagnostic zone(s), which may then be correlated to the presence or amount of the markers of interest. In preferred embodiments, one or more of the diagnostic zones comprise an antibody that binds for detection the particular marker to be detected at that particular addressable location.

[0056] The term "discrete" as used herein refers to areas of a surface that are non-contiguous. That is, two areas are discrete from one another if a border that is not part of either area completely surrounds each of the two areas.

[0057] The term "independently addressable" as used herein refers to discrete areas of a surface from which a specific signal may be obtained.

[0058] The term "antibody" as used herein refers to a peptide or polypeptide derived from, modeled after or substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof, capable of specifically binding an antigen or epitope. See, e.g. *Fundamental Immunology*, 3<sup>rd</sup> Edition, W.E. Paul, ed., Raven Press, N.Y. (1993); Wilson (1994) *J. Immunol. Methods* 175:267-273; Yarmush (1992) *J. Biochem. Biophys. Methods* 25:85-97. The term antibody includes antigen-binding portions, i.e., "antigen binding sites," (e.g., fragments, subsequences, complementarity determining regions (CDRs)) that retain capacity to bind antigen, including (i) a Fab fragment, a monovalent fragment consisting of the VL, VH, CL and CH1 domains; (ii) a F(ab')<sub>2</sub> fragment, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fd fragment consisting of the VH and CH1 domains; (iv) a Fv fragment consisting of the VL and VH domains of a single arm of an antibody, (v) a dAb fragment (Ward et al., (1989) *Nature* 341:544-546), which consists of a VH domain; and (vi) an isolated complementarity determining region (CDR). Single chain antibodies are also included by reference in the term "antibody."

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0059] In the following, the invention will be explained in further detail with reference to the drawings, in which:



[0060] Figure 1 is a chart illustrating an exemplary distribution of levels of a particular marker among a set of diseased patients and a set of non-diseased patients;

[0061] Figure 2 is a chart illustrating an exemplary scatter distribution of levels of a particular marker among a set of diseased patients and a set of non-diseased patients;

[0062] Figure 3 is an exemplary receiver operating characteristic (ROC) curve for the marker level distributions illustrated in Figure 2;

[0063] Figure 4 illustrates the chart of Figure 1 with the marker values being mapped to an indicator value;

[0064] Figure 5 is a chart illustrating an exemplary scatter distribution of panel responses for the set of diseased patients and the set of non-diseased patients;

[0065] Figure 6 illustrates a ROC curve for the panel response distributions of Figure 5 with the knee of the ROC curve labeled;

[0066] Figure 7 illustrates the progression of ROC curves through an optimization process;

[0067] Figure 8 is a chart illustrating the relative contributions of each marker in a panel;

[0068] Figure 9 shows the individual ROC curves and areas for each of 5 markers comprising the panel for figures 6 and 16;

[0069] Figure 10 shows the initial and final ROC curves for an optimization of 38 markers;

[0070] Figure 11 shows the ranking and relative average contributions of 38 markers after 50 optimizations;

[0071] Figure 12 shows the initial and final ROC curves for an optimization of 19 markers;

[0072] Figure 13 shows the ranking and relative average contributions of 19 markers after 50 optimizations;

[0073] Figure 14 shows the initial and final ROC curves for an optimization of 10 markers;

[0074] Figure 15 shows the ranking and relative average contributions of 10 markers after 50 optimizations;

[0075] Figure 16 shows the initial and final ROC curves for an optimization of 5 markers;

[0076] Figure 17 shows the ranking and relative average contributions of 5 markers after 50 optimizations;

[0077] Figure 18 shows the optimized ROC curves of 6, 3, and 2 measured and derived markers and 3 measured markers used to diagnose AMI; and

[0078] Figure 19 shows the relative contributions of all 6 of the measured and derived markers for AMI.

#### DETAILED DESCRIPTION OF THE INVENTION

[0079] In accordance with the present invention, there are provided methods and systems for the identification and use of a panel of markers for the diagnosis and/or prognosis of one or more conditions or diseases, such as cardiovascular diseases and strokes, in a subject.

[0080] Method for Defining Panels of Markers

[0081] In practice, data may be obtained from a group of subjects. The subjects may be patients who have been tested for the presence or level of certain markers. Such markers are well known to those skilled in the art. A particular set of markers may be relevant to a particular condition or disease. The method is not dependent on the actual markers. The markers discussed in this document are included only for illustration and are not intended to limit the scope of the invention. Examples of such markers and panels of markers are described in pending U.S. patent application Serial Number 10/139,086, entitled "DIAGNOSTIC MARKERS OF ACUTE CORONARY SYNDROMES AND METHODS OF USE THEREOF," and U.S. Patent Application Serial Number 10/225,082, entitled "DIAGNOSTIC MARKERS OF STROKE AND CEREBRAL

INJURY AND METHODS OF USE THEREOF," each of which is assigned to the assignee of the present application and is incorporated herein by reference. In accordance with the disclosed embodiments of the present invention, "markers" may also include derived markers as defined above, and factors such as a patient's history, sex, age and race, for example.

[0082] The group of subjects is divided into at least two sets. The first set includes subjects who have been confirmed as having a disease or, more generally, being in a first condition state. For example, this first set of patients may be those that have recently had a stroke. The confirmation of this condition state may be made through more rigorous and/or expensive testing. For purposes of this document, it will be assumed that this testing is able to confirm the condition state. Hereinafter, subjects in this first set will be referred to as "diseased".

[0083] The second set of subjects are selected from those who do not fall within the first set. This set may include all remaining subjects, or only those subjects being in a second condition state. Subjects in this second set will hereinafter be referred to as "non-diseased". Preferably, the first set and the second set each have an approximately equal number of subjects. The first and second sets of data are said to be a group of data. Multiple groups of data may be defined by repeating the steps above for different disease states, condition states, or any other selection criteria.

[0084] The data obtained from subjects in these sets includes levels of a plurality of markers. Preferably, data for the same set of markers is available for each patient. This set of markers may include all candidate markers, which may be suspected as being relevant to the detection of a particular disease or condition. Actual known relevance is not required. Embodiments of the methods and systems described herein may be used to determine which of the candidate markers are most relevant to the diagnosis of the disease or condition.

[0085] The levels of each marker in the two sets of subjects may be distributed across a broad range, as illustrated in Figure 1. Further, although Figure 1 illustrates a distribution for the marker levels of the two sets, data for the two sets may simply be available as data points for each patient, as illustrated in Figure 2. No specific distribution fit is required.

[0086] As noted above and as illustrated clearly in Figures 1 and 2, a marker often is incapable of effectively identifying a patient as either diseased or non-diseased. For example, if a patient is measured as having a marker level that falls within the overlapping region, the results of the test may not be clinically relevant.

[0087] A cutoff may be used to distinguish between a positive and a negative test result for the detection of the disease or condition. Changing the cutoff trades off between the number of false positives and the number of false negatives resulting from the use of the single marker, or in the method described herein, the panel response.

[0088] The effectiveness of a test having such an overlap is often expressed using a ROC (Receiver Operating Characteristic) curve. Other measures, such as positive predictive value (PPV) and negative predictive value (NPV) may also be used as a measure of the effectiveness of the test. ROC curves are well known to those skilled in the art. For further details, see Zweig, MH. & Campbell, C.C., *Clin Chem* 39, 561-577 (1993) and Henderson, A.R., *Ann. Clin. Biochem* 30, 521-539 (1993).

[0089] Figure 3 illustrates an example of a ROC curve for the marker level distributions of Figure 1. The horizontal axis of the curve represents (1- specificity), which increases with the rate of false positives. The vertical axis of the curve represents sensitivity, which increases with the rate of true positives. Thus, for a particular cutoff selected, the values of specificity and sensitivity may be determined. The area under the ROC curve is a measure of the utility of the measured marker level in the correct identification of one or more diseases or conditions. Thus, the area under the ROC curve can be used to determine the effectiveness of the test.

[0090] As discussed above, the measurement of the level of a single marker may have limited usefulness. The measurement of additional markers provides additional information, but the difficulty lies in properly combining the levels of two potentially unrelated measurements.

[0091] In the methods and systems according to embodiments of the present invention, data relating to levels of various markers for the sets of diseased and non-diseased patients may be used to develop a panel of markers to provide a useful panel response. The data may be provided in a database such as Microsoft Access, Oracle, other SQL databases or

simply in a data file. The database or data file may contain, for example, a patient identifier such as a name or number, the levels of the various markers present, and whether the patient is diseased or non-diseased. Thus, a chart similar to Figure 2 may be generated for each marker of interest. In practice, the generation of the chart is generally not required since the data may be directly accessible through the database or the data file.

[0092] In a preferred embodiment, one or more 'derived markers', which are a function of one or more measured markers, may be incorporated into the set of markers being studied. For example, derived markers may be related to the change in one or more measured marker values, or may be related to a ratio of two measured marker values, or may be other panel responses. In many diseases there will be a rapid change in a marker's value some time after an event. For example, following an acute myocardial infarction, (AMI), myoglobin may rise rapidly and peak about 3 hours from the event. It may then decay back to its nominal value. Looking for changes in markers can be powerful diagnostic tool. Thus, the change in myoglobin over a period of an hour, for example, may be used as a "marker" in the panel.

[0093] In practice diagnosis of a disease state from multiple markers can be confusing. Often the individual marker values may seem to contradict one another. In panels where the individual markers are not very effective, it is extremely difficult to understand their meaning. In a preferred embodiment, a function that combines the marker values into a scalar value that increases with increasing likelihood of disease is defined. In this manner, the information from multiple markers may be presented in a useable form. This defined function is referred to herein as the panel response (PR), and is a function of the marker values ( $M_{0-n}$ ), written as  $PR = f(M_{0-n})$ . The functional form of the panel response may be derived from knowledge of the pathology of the disease, or may be part of the search space. Including the functional form as part of the search space may significantly increase the number of degrees of freedom. In some cases the problem may become under specified. However, searching for the functional form may lead to further understanding of the pathology of the disease and the markers. The panel response may be scaled such that all values are between 0 and 1. Because the effectiveness of the test may not depend on a scaling of the panel response, scaling may not influence the result of the method. However forcing the panel response to be a given scale may remove an unneeded redundancy, as panel response functions that differ only by a scaling factor may in fact represent the same

solution. The panel response may also be a general function of several parameters including the marker levels and other factors including, for example, a patient's history, age, race and gender of the patient.

[0094] In a preferred embodiment, the panel response (PR) for each subject is expressed as:

$$PR = \sum_{\text{Markers}} I_i(M_i) \cdot W_i,$$

where  $i$  is the marker index,  $W_i$  is the weighting coefficient for the marker  $i$ ,  $M_i$  is the marker value for marker  $i$ ,  $I$  is an indicator function for marker  $i$ , and  $\sum$  is the summation over all candidate markers. The weighting factors scale the indicator functions and may allow for more important or specific markers to have a greater impact on the final panel response. The indicator function maps the marker value into a functional form appropriate to the marker's pathology. The indicator functions can be complex and should be chosen to match the marker. This will be illustrated in the embodiments described below. The indicator function may be a different functional form for each marker. In one example, the indicator function can map the marker value into a probability of the disease state. This mapping may not be a simple function of the marker value. In this example the said indicator from each marker can be summed to determine a relative index which is related to the probability of the patient being diseased. In a preferred embodiment the sum of all the weighting coefficients is constrained to a particular value, such as 1.0. As described above, this removes redundancy but doesn't change the objective function. In a preferred embodiment the indicator function is constrained to values between 0 and 1. In a further preferred embodiment, both of the above constraints are satisfied, thus, the panel response is also constrained to a value between 0 and 1.

[0095] In many disease states such as stroke, nonspecific markers associated with that state are elevated. But above a certain threshold, higher values of the marker may not relate to a higher probability of disease state. Below a certain threshold, lower marker values may not relate to a lower probability of disease state. In this situation the indicator function may not increase linearly with the marker value. A preferred embodiment is an indicator function that is a function that has a high and monotonic rate of change between the thresholds, and a small rate of change elsewhere. Examples of this type of function are

the ramp, step, or sigmoid functions. One may associate the lower threshold with the start of an overlap region (or cutoff region), and the upper threshold with the end of the overlap region, as shown in Figure 4. Below the lower threshold the probability of disease is substantially 0, while above the upper threshold the probability of disease is 1. Note that in the case where the indicator function is a step function and the weighting value is 1 for each marker, then the panel response is simply the number of markers above the cutoff. This case is identical to the example used above where one is searching for the best panel with  $n$  of  $m$  markers above their cutoff. Allowing the indicator to vary continuously near the threshold enables the panel response to be sensitive to a marker just under the cutoff. This information is not lost as it is in the  $n$  of  $m$  marker example or the step function example, where the indicator value is not continuous. Another common approach of summing over  $M \cdot W$  forces the linear relation with the marker value,  $M$ . But as discussed above the most appropriate indicator function may not increase linearly with the marker value. In a further preferred embodiment the ramp function is used as an elevation indicator function. As illustrated in Figure 4, the indicator values between the threshold regions may vary linearly from a value of zero at one end to a value of one at the other end. In other embodiments, non-linear variations of the indicator value may be used. The ramp function has the advantage of simplicity, and may be good approximation to other function in this class. With proper choices of parameters, the ramp function can be equivalent to the step function or can increase linearly with the marker value.

[0096] In some disease states, for example unstable angina, a specific marker such as the cardiac troponins (including isoforms of cardiac troponin, comprising troponin I and T and complexes of troponin I, T and C) may be elevated above the normal population, but further elevation indicates an acute condition, in this case a myocardial infarction.

Unstable angina is an ischemic condition that leads to minor necrosis of cardiac tissue. During a myocardial infarction, there is major necrosis of cardiac tissue. Cardiac troponin, which is specific to cardiac necrosis, is elevated in both conditions, but the amount of elevation is related to the amount of necrosis. The best indicator function of cardiac troponin in diagnosing unstable angina may not be an elevation indicator function. In a preferred embodiment the indicator function may be a function that is peaked near the expected values of unstable angina, and decreases when the marker value is above or below the expected value. Examples of this type of function include a Gaussian, triangle, trapezoid, or square function. These functions tend to localize the marker value of interest

around a specific value. Another example of use for such an indicator function is in cases where a pattern of markers values indicates a disease state. For example, a disease condition may be indicated when one or more markers are within a range of values. When desired, the use of this type of indicator may allow for recognition of patterns of marker values.

[0097] It is possible that one of the markers in the panel is specific to the disease or condition being diagnosed. An example of such a marker is cardiac specific cardiac troponin when used in the diagnosis of acute myocardial infarction. The role of TnI is described above. The panel response can be coded for markers that are specific, and the information may be used during the optimization of the panel response parameters. Typically the specific cutoff of such markers is known, so the specific cutoff values may not be included as a search parameter. When such a marker is present at above or below a certain threshold (e.g., the specific cutoff), the panel response may be set to return a "positive" test result, regardless of the levels of non-specific markers. When the specific cutoff is not satisfied, however, the level of the specific marker may nevertheless be used as a possible contributor to the panel response, along with the remaining markers on the panel.

[0098] In an example where the panel is being chosen based on  $n$  of  $m$  markers being elevated, the effectiveness of the panel is dependent on the choice of  $n$ . This extra dimensionality can be eliminated by using an objective function. The reduction of dimensionality may simplify the search process, and the objective function provides a scalar value that is optimized during the search process. The objective function should generally be indicative of the effectiveness of the panel, as may be expressed by, for example, overlap of the panel responses of the diseased set of subjects and the panel responses of the non-diseased set of subjects. In this manner, the objective function may be optimized to maximize the effectiveness of the panel by, for example, minimizing the overlap. In a preferred embodiment, the ROC curve representing the panel responses of the sets of subjects may be used to define the objective function. A ROC curve with a high value for the ROC curve area indicates a test with a good ability to discriminate between diseased and non-diseased. So, continuing with the  $n$  of  $m$  example above, there should exist a value of  $n$  which yields a clinically relevant test, but the value of  $n$  need not be determined during the search process. The objective function is the scalar response that is



maximized by the search algorithm. The objective function can include the correlation to a quantitative measure of disease, for example, the NIH score for stroke patients. Other measures of effectiveness may include, for example, odds ratios, a positive predictive value (PPV) and a negative predictive value (NPV) of the panel. The odds ratio, PPV and NPV are well known to those skilled in the art. One skilled in the art will recognize there other measures of the effectiveness of the test. See *The Immunoassay Handbook*, Second Edition, David Wild, 2001 for measures of effectiveness. Many common measures of effectiveness require the selection of a cutoff value. These functions may still be used, and the cutoff value may also be included as a search parameter. In a preferred embodiment objective functions are chosen that do not require the selection of a cutoff value. The measure that is most appropriate for defining an effective test may vary.

[0099] In a preferred embodiment, the area under the ROC curve representing the panel responses of the sets of subjects may be used to define the objective function. Those skilled in the art will recognize that the area of the ROC curve is a measure of the effectiveness of the test. An area of 1 corresponds to a perfect test, and an area of 0.5 corresponds to a random test.

[0100] In another embodiment, the knee of the ROC curve is used for the objective function. The knee of the ROC curve is the point illustrated in Figure 6, and the value is represented as the product of the specificity and sensitivity at the knee. In one embodiment the knee is found by maximizing the product of Specificity and Sensitivity. Higher knee values may indicate squarer ROC curves.

[0101] In another embodiment the objective function is the specificity at a prescribed sensitivity. If one requires that a test have only a certain sensitivity (ability to detect diseased patients) then maximizing the specificity, which may reduce the number of false positives, may improve the clinical effectiveness of the test.

[0102] In another embodiment the objective function is the sensitivity at a prescribed specificity. If one requires that a test have only a certain specificity (the number of false positives), then maximizing the sensitivity, which may increase the ability to detect diseased patients, may improve the clinical effectiveness of the test.

[0103] The objective function is a function of two or more of these characteristics. In a preferred embodiment, the objective function is the product of two or more characteristics of the ROC curve. An example of this is to use the product of the ROC curve area, knee, sensitivity at a prescribed specificity, and specificity at a prescribed sensitivity. Any one characteristic alone may not result in a desired solution. By using the product of two or more of these, a more desirable solution may be achieved.

[0104] Variations in the values of markers over some time interval within a patient may be a powerful tool in the diagnosis of disease states or condition or the progression of disease states or conditions. As defined above the panel response can be thought of as a new derived marker, where the panel response value is thought of as the marker value. Changes in the panel response value over some time interval within a patient may be a powerful tool in the diagnosis of disease states or conditions or the progression of disease states or conditions. The change in the panel response can also be used as a derived marker. One can apply all of the ideas and methods discussed in this document to the case where a derived marker is a panel response or responses or the change in the panel response or responses. Calling the change in the panel response a derived marker may be equivalent to defining a new panel response that is the change in the panel response over some time interval. The new panel response function is a function of the marker values at two time points. All methods and ideas discussed in this document can apply to the new panel response.

[0105] The two subject populations ("diseased" and "non-diseased") may contain subgroups. Thus, in accordance with the methods described herein, panel responses may be defined between such subgroups, considering one subgroup as a new "diseased" population and a second as a new "non-diseased" population. Accordingly, a panel response combining the new panel responses for these new "diseased" and "non-diseased" populations can be defined. For example, in a stroke data set, an original "non-diseased" population may contain both normals and stroke "mimics." Similarly, the "diseased" population may contain both ischemic and hemorrhagic stroke. Panel responses can be defined for normal vs. mimic, normal vs. ischemic, normal vs. hemorrhagic, mimic vs. ischemic, mimic vs. hemorrhagic, and ischemic vs. hemorrhagic. Once these panel responses are defined, a panel response using these five panel responses for inputs either

alone or in conjunction with the other markers can be defined to distinguish non-diseased from diseased.

[0106] Searching for the best panel can be accomplished by trying all the different combinations of parameters of the panel response function. But with panels of 40 markers, and just one degree of freedom per marker, taking 10% steps in the parameter values will require  $10^{40}$  iterations. The age of the universe is estimated to be about 20 billion years or about  $6.3 \times 10^{17}$  seconds. Clearly this approach is not practical, and the problem requires the use of a search engine. Optimization algorithms are well-known to those skilled in the art and include several commonly available minimizing or maximizing functions including the Simplex method and other constrained optimization techniques. It is understood by those skilled in the art that some minimization functions are better than others at searching for global minimums, rather than local minimums. Many of these exist, and detailed descriptions can be found in the literature. For more information on minimization and maximization functions, reference may be made to Numerical Recipes in C, The Art of Scientific Computing, Second Edition, W. Press, et al., Cambridge University Press, 1992, which is hereby incorporated by reference. The panel response and the objective function have helped enable the use of search routines. The objective function value is the response that the search routine will maximize, and the parameters of the panel response function form the n-dimensional space to be searched. While the objective function does not need to be continuous, i.e. it may have discrete values, panel response functions that are continuous may reduce the granularity of the objective function. This may help the algorithm find better solutions. While many search routines will in fact look for minima, the problem may be inverted by minimizing  $(-1) \times \text{Objective Function}$ .

[0107] In a preferred embodiment the search engine uses the Downhill Simplex Method in Multidimensions. This method is described in Numerical Recipes in C, The Art of Scientific Computing, Second Edition, W. Press, et al., Cambridge University Press, 1992. The simplex has  $n+1$  vertices, where  $n$  is the number of dimensions or degrees of freedom. The routine 'walks' the simplex along the  $n$  dimensional surface, moving one vertex at a time. The scale of the simplex can change so it can both quickly walk in downhill directions and crawl through tight crevices. The routine may not find a global minimum because it can become trapped in a local minimum. The simplex will search all real space. The parameters of the panel response are often valid only within some range,

defining the bounds of the system. The simplex must be constrained to only search in this space, and there must be no degeneracy introduced when approaching such a constraint. One skilled in the art will recognize that there are many ways to address this constraint. In a preferred embodiment a penalty is assessed when a vertex moves out of bounds. This penalty creates steep canyon walls around the bounds of the system, effectively constraining the simplex within the bounds of the system.

[0108] A well-known limitation of search engines is their tendency to find only a local minimum, typically not the global minimum. Several techniques are known to improve the ability to seek out the global minimum. In a preferred embodiment, the technique of simulated annealing is used. This method is also described in Numerical Recipes in C. The Art of Scientific Computing, Second Edition, W. Press, et al., Cambridge University Press, 1992. Simulated annealing adds a random error to each decision of the search engine. This random error gives the search engine the ability to move out of a shallow local minimum, so it can seek out a deeper solution. The random error is systematically reduced until a minimum is found. The random error is similar to the effect of temperature in annealing processes. The scale of the random error is said to be the temperature. The annealing process may improve the chances of finding a global, rather than local, minimum. The annealing process may result in a more stable solution since the random variation may move the simplex out of a narrow, unstable region. The optimization process may be terminated when the difference in the objective function between two consecutive iterations is below a predetermined threshold, thereby indicating that the optimization algorithm has reached a region of a local minimum. The number of iterations may also be limited in the optimization process.

[0109] The selection of the initial conditions, for example the initial simplex value, may affect the optimization process. So, generally good selections of the initial parameters are sought. In the example of a search using a simplex, all vertices of the simplex must be initialized. If only one good vertex is defined, the other vertices can be assigned by applying a random deviation to each parameter. The scale of this random deviation sets the scale of the initial simplex. For example when elevation indicator functions are used, the location of the cutoff region may initially be selected at any point. But, selection near a suspected optimal location may facilitate faster convergence of the optimizer. In a preferred method, the cutoff region is initially centered about the center of the overlap

region of the sets of patients. In one embodiment, the cutoff region may simply be a cutoff point. In other embodiments, the cutoff region may have a length of greater than zero. In this regard, the cutoff region may be defined by a center value and a magnitude of length. In practice, the initial selection of the limits of the cutoff region may be determined according to a pre-selected percentile of each set of subjects. For example, a point above which a pre-selected percentile of diseased patients are measured may be used as the right (upper) end of the cutoff region. In another embodiment the weighting factors may initially be all set to one. In a preferred embodiment, the initial weighting coefficient for each marker may be associated with the effectiveness of that marker by itself. For example, a ROC curve may be generated for the single marker, and the area under the ROC curve may be used as the initial weighting coefficient for that marker. This gives more weight to markers with better univariate utility.

[0110] Having selected optimal parameters for the panel response function, the panel responses for each subject in each set of subjects, and the distribution of the panel responses for each set may now be analyzed. Figure 9 shows the ROC curves and area of several markers that have a poor diagnostic utility. The markers data are used to generate Figure 5. When the poor markers are combined and the panel response determined, the results show that the panel now has enhanced utility. Figure 5 illustrates an exemplary distribution of the panel responses for diseased and non-diseased subjects. Based on these distributions, a ROC curve may be generated, as illustrated in Figure 6. The ROC curve illustrated in Figure 6 reflects optimized values for the weighting coefficients and the thresholds for a ramp indicator function.

[0111] Figure 7 illustrates an exemplary progression of a ROC curve through a plurality of iterations of an optimization process in which the objective function is defined as the area under the ROC curve. As illustrated in Figure 7, as the number of iterations increases, the area under the curve may progressively increase. Thus, the optimization process may provide a panel response function for the markers. In this example, the indicator function is a ramp function. The optimization routine found values of the weighting coefficients and high and low threshold values which are represented as a cutoff value and linear range. Table 1 illustrates a panel of 38 candidate markers with weighting coefficients and cutoff regions resulting from the optimization process. The 38 markers are listed generically as Analyte 1 through Analyte 38. The sense of each marker, as

described above, is also indicated in Table 1, with "Incr" representing a positive sense and "Decr" representing a negative sense. The cutoff location indicated in Table 1 refers to the marker level value around which the cutoff region is centered, while the length of the cutoff indicates the range of marker level values covered by the cutoff region. In this manner, any number of markers may be used to develop a highly effective panel response function that can be used for the diagnosis of a disease or condition.

[0112] The result of any given search is likely not to be the global minimum. It may be any local minimum that the search engine settled in. In a product to be used for clinic diagnosis, it is preferable to find a very stable solution. Inaccuracy associated with the measurement of the marker values should not significantly influence the effectiveness of the test. Also, the defining data may not be inclusive of all patients; it may be only a small sample, and the remaining population may deviate from the defining population. The desired characteristics of the minimum may include a wide width and shallow side walls. In a three-dimensional analogy, we would prefer a minimum like a crater as opposed to a mine shaft. One method to seek out these types of solutions is to search multiple times. If a statistically significant number of optimizations is performed, then the better solutions will be the largest group of similar results. This is because, using the example above, it may be more likely to find the crater than the mine shaft.

[0113] As discussed above, not every minimum found may be desirable to use. Generally stable parameters are desired, meaning that variations in the marker values or parameters do not adversely impact the effectiveness of the test. The width and depth of the minimum may provide an indication of the stability of the solution. In addition, one could use bootstrapping methods known in the art to determine the stability of a particular solution. In such methods, a "training" data set is used to arrive at an initial solution. That solution is then validated by applying the solution to a second "validation" data set that is independent of the training data set.

[0114] There are several additional examples of methods that may quantify the quality of a set of parameters, e.g., when a validation data set is unavailable. A first example is to vary the marker values by some random percentage. By doing this one can simulate all the variations expected due to assay imprecision, biological variations, and any other source of uncertainty. For example, variations in marker values may relate to the relative

imprecision of the test that was used to generate the data. One skilled in the art will recognize that there are limits to the analytical precision of a test. For example, in the immunoassay art, it is common to encounter 5-20% coefficients of variations of the tests. Therefore, when considering the imprecision of the testing methodology, the parameters remain stable relative to the imprecision of the methodology. The randomized data set can be reanalyzed to generate the new panel response ROC curve and objective function value. An acceptable deterioration may indicate the parameters give a solution that is stable to variations in marker values and may also verify that the solution does not simply fit the noise in the data.

[0115] A second example would be to vary one or more of the parameters in the panel definition by some amount. The change in the objective function value may be a measure of the quality of the solution. Each parameter could be varied independently to determine the stability of each parameter. Or all parameters may be varied randomly to sample the space surrounding the solution. The magnitude of the variation can be constrained to sample a certain volume or shell surrounding the solution. Thus the deterioration of the objective function as a function of displacement from the solution may be determined.

[0116] In a third example, a seed simplex is generated with a given length scale about the known minimum. The length scale of the seed simplex can be systematically increased until re-optimizations lead to a different minimum, i.e. the solution is no longer recovered. The length scale, which results in finding new minimums, may be related to the width of the minimum. In a fourth example, using the final simplex of the optimization, the temperature can be systematically increased until re-optimizations lead to a different minimum, i.e. the solution is no longer recovered. The temperature, which results in finding new minimums, may be a measure of the depth of the solution. In a fifth example, most common solutions from the multitude of optimizations, may represent the most stable solution. The common solutions can be grouped based on their similarity. Correlation techniques and clustering techniques can be used to group the solutions, and are well known to one skilled in the art. Solutions can also be grouped by finding ones that, using a three dimensional analogy, are on the same valley floor. These can be grouped by calculating the reduction in objective function while moving from one solution to another. In the analogy above this determines how high of a hill lies between the two solutions. From the teaching above, it is now clear that other approaches exist for

quantifying the quality of a set of parameters, and the examples above are not intended to limit the invention.

[0117] The use of the term “non-diseased” does not mean that the particular subject is disease-free, only that the subject is free from the one or more diseases or conditions being evaluated. In practice, a pre-filtering of subjects may be performed on the basis of any particular characteristic of the subjects, including the existence of other diseases. For example, the method and systems described may be applied to first divide a group of subjects into “diseased” and “non-diseased” for Disease A, and then divide the group into “diseased” and “non-diseased” for Disease B. A panel of markers for each disease may then be determined. In another embodiment, the same panel of markers may be used for both diseases with a different set of parameters, such as weighting coefficients, for each disease. In another embodiment subjects with disease A can be defined as non-diseased, and subjects with disease B can be defined as diseased. In this embodiment the described techniques can be employed to determine a panel that differentiates between diseases A and B.

[0118] The search routine will optimize the objective function or functions selected on the specified data set. But often times it is important to constrain or optimize a second group of data simultaneously. This is accomplished by pre-filtering the source data to get the two or more groups of data of interest. Different objective functions can be selected for each group of data, and the search engine can find the minimum of the product of objective functions. The objective function of one of the groups of data can also be constrained to be at least some value. When the objective function is greater than or equal to this constraint, the value returned to the search engine is the constraint value. When the objective function is below the constraint value the objective function value is returned. The search routine will look for solutions that satisfy the constraint condition, but the best solution may fall outside the constraint condition. The iterations of the optimization algorithm generally vary the independent parameters to satisfy the constraints while maximizing the objective function. An example of this usage is stroke data that contains norm health donors and stroke mimics. We would like to find a panel response function that will distinguish stroke from stroke mimics, but that will also have a low false positive rate for normal healthy donors (NHD). Since the number in each sample set is not equal, simply combining the data and analyzing will not give a satisfactory result. Results will be



skewed to the data set with larger  $n$ , in our case NHD. However, if the objective functions of the two groups of data are individually calculated and combined, then the groups of data are given equal weight. In another example we want to ensure that patients presenting soon after the onset of symptoms will be properly diagnosed, but we still want to ensure that patients presenting at longer times are also properly diagnosed. Again, the population numbers will be different. So, to give equal weighting, they need to be simultaneously analyzed as two groups of data. Other constraints may include limitations on one group of samples while optimizing for an objective function for a second group. For example, a panel may be optimized for one disease while the same panel may be constrained to provide at least an acceptable minimum value for the area under a ROC curve for a second disease.

[0119] Within the teachings of this document we have used for simplicity markers that are elevated in patients with the disease or positive sense markers. However this is not always the case, and often, particularly with poor univariate markers, it is not clear from univariate analysis whether the marker when used in conjunction with the other markers in the panel, is best utilized in a positive or negative sense. If the sense of a marker is inverted, then it is straightforward to invert the indicator function for that marker. If the sense is not known, then the search engine may include this as a degree of freedom. For example, in one embodiment, the sense may be a truly separate independent variable, which may be flipped between positive and negative by the optimization process. For optimal performance, the sense should map smoothly from improper to proper, and there should be pressure (or a gradient) that allows the search engine to move toward the proper sense. In a preferred embodiment the sense is switched by allowing the weighting coefficient of the analyte to go negative. If the wrong sense is selected, the weighting coefficient will be driven towards zero since inclusion of the marker in the panel response negatively impacts the objective function. The search engine will be able to drive the weighting coefficient across zero to the proper sense. In this example, the negative weight is just a flag to invert the sense. The absolute value of the weight is used in the panel response function. But this allows for a continuous function moving from positive to negative sense.

[0120] In order to determine the best panel, which for practical reasons may often mean 10 or less markers, one must find a way to systematically remove markers that do

not significantly contribute to the overall result. Again, there are several methods that may be applied to arrive at a useful set of markers. For example, markers may be initially selected by performing univariate statistics to determine if a marker provides a meaningful distinction between the "disease" and "non-disease" groups. While the distinction in such univariate methods may be weak, any correlation to the disease or non-disease state may be enough to indicate that a marker should be considered for further analysis.

[0121] One may also calculate the contribution from each marker to the marker panel result itself. A method to accomplish this is to remove an analyte from the panel, and recalculate the objective function. This can be achieved by re-optimizing the parameters in the absence of the analyte and determining the best objective function. One can also remove the analyte and recalculate the objective function without re-optimizing. While not as precise, it can offer significant savings computationally. In either case, care must be taken to ensure that the data set is static when the analyte is removed. The change in the objective function is related to the contribution of the marker. This method for identifying the relative importance of each marker is illustrated in Figure 8. The resulting changes in the objective function are noted for each marker and plotted, as shown in Figure 8. Figure 8 illustrates the effect each marker has on the various features of the ROC curve corresponding to the panel responses for the two sets of subjects. The various ROC-curve features noted in Figure 8 include the area under the ROC curve, the location of the knee of the ROC curve, the sensitivity at a predetermined specificity, and the specificity at a predetermined sensitivity. The markers may then be arranged in order of decreasing contribution, as illustrated in Figure 8. The vertical axis in Figure 8 indicates the relative change in the values of the various ROC-curve features. In embodiments where a weighting coefficient is applied to each analyte, the weight for the analyte can be set to zero to remove the analyte's contribution to the panel result. The change in objective function can then be determined. In embodiments where a weighting coefficient is applied to each analyte, one can not simply use the weights as the contribution. An example of why this does not give the proper result is the case where a marker has zero impact on the test. In this case, the weight it is given by the search program can be any value, so it is possible that its weight will be the highest.

[0122] In order to develop lower-cost panels, which require the measurement of fewer marker levels, certain markers may be eliminated from the panel. In this regard, the

effective contribution of each marker in the panel may be determined to identify the relative importance of the markers. Once the relative contributions are calculated then one can rank them from largest to lowest. The markers with the largest changes in objective function may be the ones with most contribution. The ones with the least change in objective function may be the ones with the least contribution. However, for example, if two markers are perfectly correlated, then the combined contribution from both may be equivalent to the contribution of just one if the second one is removed. The partitioning of the contributions is not necessarily equal. So an important marker may not have a high contribution. This problem can be avoided by first looking at the correlation, or "interaction," between markers, or by removing only one marker or more with the lowest contribution. Methods for calculating interaction terms are well known to those of skill in the art.

[0123] From the discussion above, it is noted that it may not be prudent to just select the top 3 markers from a panel of 40. Depending on the number of target markers being searched and the size of the target panel, one may want to eliminate only the marker with the lowest contribution or the lowest markers, and repeat the process until the target panel size is reached. With properly defined panel responses, markers of no importance may not adversely impact the objective function. This is because a) the search routine may chose parameters such that the marker is not used, and b) in general a random marker will not change the objective function. So, starting with a large panel and reducing it to the desired size will lead to the optimum panel. But the objective function may degrade as markers are eliminated. One may have to trade off panel effectiveness with the number of markers. For example, in order to obtain a panel of ten markers, the ten highest-rated markers, i.e. those on the left side in Figure 8, may be selected. For example, Analytes 38, 1, 16, 33, 27, 12 and 8 may be selected in a final panel of markers. In a preferred embodiment, only a few of the markers on the right side may be eliminated, and the remaining markers in the panel may be optimized. For example, Analytes 31, 24, 25, 4 and 10 may be eliminated in a first round, and the optimization and ranking procedures may be repeated with the remaining 33 markers. This results in a chart similar to that shown in Figure 8, but with fewer markers. This process may be repeated until a desired number of markers remains in the panel.

[0124] It is possible that one of the markers in the panel is specific to the disease or condition being diagnosed. An example of such a marker is cardiac specific TnI when used in the diagnosis of acute myocardial infarction. The role of TnI is described above. The panel response can be coded for markers that are specific, and the information is used during the optimization of the panel response parameters. Typically the cutoff of such markers is known, so the cutoff values may not be included as a search parameter. When such a marker is present at above or below a certain threshold, the panel response may be set to return a "positive" test result, regardless of the levels of non-specific markers. When the threshold is not satisfied, however, the level of the specific marker may nevertheless be used as possible contributor to the objective function, along with the remaining markers on the panel.

[0125] In a preferred embodiment the panel will include markers derived from the rate of change of markers measured by the panel. In a further preferred embodiment the panel will have two panel response functions, one that utilizes the derived markers when present, and when not present one that does not utilize the derived markers. The two panel response functions may use different parameters. These parameters may be obtained by optimizing the data with and without utilizing the derived marker or markers. For example, a patient may be measured when first arriving at the hospital for a particular set of markers. Since there is only one sample time for the patient a panel response function which does not include marker changes is used. The patient would be diagnosed as diseased or non-diseased based on the results of the test. The same patient may be measured again an hour later. Now there are two points, and so a second panel response function which utilizes marker changes is used. The use of this response function is important when a marker or panel of markers of disease indicates non-disease, but the change (usually increase) in the value of one or more markers represents the onset of disease.

[0126] It is possible for a panel of markers to contain enough information to diagnose a multitude of conditions. In the simplest case, the markers used in the diagnosis of condition A are different from the markers used in the diagnosis of condition B, but the panel contains the union of the markers for A and B. In a preferred embodiment, the markers used in the diagnosis of condition A contains at least one of the markers used in

the diagnosis of condition B. In a future preferred embodiment there is a high degree of overlap in the markers used to diagnose a multitude of conditions.

[0127] The method described above may be implemented in a variety of manners. In a preferred embodiment, the method is implemented as a program product, such as a software package. The program product may be implemented on a computer, such as a personal computer, a mainframe or a handheld device. It will be apparent to those skilled in the art that the program product may be implemented on a device in any number of ways including software, firmware, etc. In one embodiment, the program product is implemented on a meter which may be capable of directly measuring levels of one or more markers. For example, the program product may be implemented on a fluorometer or a reflectometer. Such devices are well known to those skilled in the art.

[0128] In a most preferred mode, patient types, disease types, and time frames are selected to provide two data sets, "diseased" and "non-diseased," which have the characteristics to be evaluated. Multiple groups of data can be selected, each set consisting of a set of diseased and as a set of non-diseased samples. The values for any derived marker values of interest are calculated for each record in the selected groups of data. This may include calculating the change in marker value or panel response value from the initial value. Based on the disease and marker pathology, a functional type for the indicator function is chosen for each marker to be included in the panel. The teachings in this document should enable one skilled in the art of the disease and marker to make the appropriate choice. If the disease or marker pathology is not sufficiently understood to choose a functional form then the functional form can become part of the search. Once the indicator functions have been defined, then the initial parameters are chosen from the univariate marker analysis. These initial parameters define one vertex of the initial simplex. The number of vertexes constituting the simplex is the number of search parameters in the panel response plus one. Each remaining vertex is populated by varying each parameter by a random amount. The scale of this random amount can be fixed to be a percentage of the parameter value. This spreads the simplex out around the initial point, and gives the simplex a size scale. The objective function for each group of data is defined by selecting any combination of the ROC curve area, the ROC knee, the ROC sensitivity, and the ROC specificity, but typically all four are selected. The objective function for each group of data can be chosen to be optimized or to maintain a minimum target value. Thus

the optimization of one group of data can be constrained such that a second group of data has at least a minimum objective function value. The parameters are then optimized to maximize the chosen objective function utilizing the downhill simplex method with simulated annealing. At the end of the optimization the relative contribution for each marker is calculated by setting the weight of that marker to zero and recalculating the panel ROC curve. When the analyte is so removed from the panel response, the new ROC curve is calculated with the identical data and no other parameters in the panel response are changed. The process of optimizing and calculating marker contributions is repeated  $n$  (~100) times. After  $n$  optimizations, the average contribution of each marker over the  $n$  optimizations is calculated, and the markers are ranked based on its average contribution. The poorest markers, typically the poorest half or less, are removed from the panel and the entire process is repeated as many times as required to reduce the panel to the desired size.

[0129] Candidate solutions are then tested for stability and robustness by varying the input data and the parameters, and finding solutions that are grouped together. The best solution may be derived from an average of one or more good solutions. Using optimal analytes and parameters for the panel response function found via the search method described above, the ROC curve of the panel response from clinical data is calculated. Based upon the panel response ROC curve an appropriate cutoff is chosen. The choice may be influenced by factors such as clinical factors, treatment methods, and cost considerations, which one skilled in the art will recognize. The panel response is calculated from the measured marker values of the patient for whom it is desired to determine the presence or absence of the target disease. Using the chosen cutoff, assign a diagnosis for the patient.

[0130] Using optimal analytes and parameters for the panel response function found via the search method described above, for panel response functions which include and exclude markers derived from the change in a measured marker, the ROC curve of the panel response from clinical data is calculated. Based upon the panel response ROC curves appropriate cutoffs are chosen for each. The choice may be influenced by factors such as clinical factors, treatment methods, and cost considerations, which one skilled in the art will recognize. Upon measurement of the initial sample, the panel response is calculated from the measured marker values of the patient for whom it is desired to determine the presence or absence of the target disease. Using the chosen cutoff, assign a

diagnosis for the patient. A second or more measurement may be required to further clarify the diagnosis. At the appropriate time interval, draw more sample from the patient and measure the marker values. Using the panel response function that includes derived markers, calculate the panel response value and determine a diagnosis by comparing the panel response value to the chosen cutoff value. The panel response of the first measurement can also be compared to panel responses determined from subsequent measurements. One skilled in the art will recognize that serial blood draws can yield critical information of the presence and progression of diseases, particularly acute diseases. If more measurements are required for proper patient treatment, continue taking samples at the desired intervals.

[0131] Measures of test accuracy may be obtained as described in Fischer *et al.*, *Intensive Care Med.* 29: 1043-51, 2003; Zhou *et al.*, *Statistical Methods in Diagnostic Medicine*, John Wiley & Sons, 2002; and Motulsky, *Intuitive Biostatistics*, Oxford University Press, 1995; and other publications well known to those of skill in the art, and used to determine the effectiveness of a given marker or panel of markers. These measures include sensitivity and specificity, predictive values, likelihood ratios, diagnostic odds ratios, hazard ratios, and ROC curve areas. As discussed above, suitable tests may exhibit one or more of the following results on these various measures:

[0132] A ROC curve area of greater than about 0.5, more preferably greater than about 0.7, still more preferably greater than about 0.8, even more preferably greater than about 0.85, and most preferably greater than about 0.9;

a positive or negative likelihood ratio of at least about 1.1 or more or about 0.91 or less, more preferably at least about 1.25 or more or about 0.8 or less, still more preferably at least about 1.5 or more or about 0.67 or less, even more preferably at least about 2 or more or about 0.5 or less, and most preferably at least about 2.5 or more or about 0.4 or less;

an odds ratio of at least about 2 or more or about 0.5 or less, more preferably at least about 3 or more or about 0.33 or less, still more preferably at least about 4 or more or about 0.25 or less, even more preferably at least about 5 or more or about 0.2 or less, and most preferably at least about 10 or more or about 0.1 or less; and/or

a hazard ratio of at least about 1.1 or more or about 0.91 or less, more preferably at least about 1.25 or more or about 0.8 or less, still more preferably at least about 1.5 or more or about 0.67 or less, even more preferably at least about 2 or more or about 0.5 or less, and most preferably at least about 2.5 or more or about 0.4 or less.

[0133] Measures of diagnostic accuracy such as those discussed above are often reported together with confidence intervals or p values. These may be calculated by methods well known in the art. *See, e.g.*, Dowdy and Wearden, *Statistics for Research*, John Wiley & Sons, New York, 1983. Preferred confidence intervals of the invention are 90%, 95%, 97.5%, 98%, 99%, 99.5%, 99.9% and 99.99%, while preferred p values are 0.1, 0.05, 0.025, 0.02, 0.01, 0.005, 0.001, and 0.0001.

[0134] Exemplary Symptom-Based Marker Panels

[0135] Patients presenting for medical treatment often exhibit one or a few primary observable changes in bodily characteristics or functions that are indicative of disease. Often, these "symptoms" are nonspecific, in that a number of potential diseases can present the same observable symptom or symptoms. A typical list of nonspecific symptoms might include one or more of the following: shortness of breath (or dyspnea), chest pain, fever, dizziness, and headache. These symptoms can be quite common, and the number of diseases that must be considered by the clinician can be astoundingly broad.

[0136] Taking shortness of breath (referred to clinically as "dyspnea") as an example, this symptom considered in isolation may be indicative of conditions as diverse as asthma, chronic obstructive pulmonary disease ("COPD"), tracheal stenosis, obstructive endobronchial tumor, pulmonary fibrosis, pneumoconiosis, lymphangitic carcinomatosis, kyphoscoliosis, pleural effusion, amyotrophic lateral sclerosis, congestive heart failure, coronary artery disease, myocardial infarction, cardiomyopathy, valvular dysfunction, left ventricle hypertrophy, pericarditis, arrhythmia, pulmonary embolism, metabolic acidosis, chronic bronchitis, pneumonia, anxiety, sepsis, aneurismic dissection, *etc.* *See, e.g.*, *Kelley's Textbook of Internal Medicine*, 4<sup>th</sup> Ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2000, pp. 2349-2354, "Approach to the Patient With Dyspnea"; Mulrow *et al.*, *J. Gen. Int. Med.* 8: 383-92 (1993).



[0137] Similarly, chest pain, when considered in isolation, may be indicative of stable angina, unstable angina, myocardial infarction, musculoskeletal injury, cholecystitis, gastroesophageal reflux, pulmonary embolism, pericarditis, aortic dissection, pneumonia, anxiety, *etc.* Moreover, the classification of chest pain as stable or unstable angina (or even mild myocardial infarction) in cases other than definitive myocardial infarction is completely subjective. The diagnosis, and in this case the distinction, is made not by angiography, which may quantify the degree of arterial occlusion, but rather by a physician's interpretation of clinical symptoms.

[0138] Differential diagnosis refers to methods for diagnosing the particular disease(s) underlying the symptoms in a particular subject, based on a comparison of the characteristic features observable from the subject to the characteristic features of those potential diseases. Depending on the breadth of diseases that must be considered in the differential diagnosis, the types and number of tests that must be ordered by a clinician can be quite large. In the case of dyspnea for example, the clinician may order tests from a group that includes radiography, electrocardiogram, exercise treadmill testing, blood chemistry analysis, echocardiography, bronchoprovocation testing, spirometry, pulse oximetry, esophageal pH monitoring, laryngoscopy, computed tomography, histology, cytology, magnetic resonance imaging, *etc.* See, e.g., Morgan and Hodge, *Am. Fam. Physician* 57: 711-16 (1998). The clinician must then integrate information obtained from a battery of tests, leading to a clinical diagnosis that most closely represents the range of symptoms and/or diagnostic test results obtained for the subject.

[0139] A first step in the identification of suitable markers for symptom-based differential diagnosis requires a consideration of the possible diagnoses that may be causative of the non-specific symptom observed. Taking dyspnea as an example, the potential causes are myriad. The following discussion considers three potential diagnoses: congestive heart failure, pulmonary embolism, and myocardial infarction; and three potential markers for inclusion in a differential diagnosis panel for these potential diagnoses: BNP, D-dimer, and cardiac troponin.

[0140] *BNP*

[0141] B-type natriuretic peptide (BNP), also called brain-type natriuretic peptide is a 32 amino acid, 4 kDa peptide that is involved in the natriuresis system to regulate blood

pressure and fluid balance. Bonow, R.O., *Circulation* 93:1946-1950 (1996). The precursor to BNP is synthesized as a 108-amino acid molecule, referred to as "pre pro BNP," that is proteolytically processed into a 76-amino acid N-terminal peptide (amino acids 1-76), referred to as "NT pro BNP" and the 32-amino acid mature hormone, referred to as BNP or BNP 32 (amino acids 77-108). It has been suggested that each of these species -- NT pro-BNP, BNP-32, and the pre pro BNP -- can circulate in human plasma. Tateyama *et al.*, *Biochem. Biophys. Res. Commun.* 185: 760-7 (1992); Hunt *et al.*, *Biochem. Biophys. Res. Commun.* 214: 1175-83 (1995). The 2 forms, pre pro BNP and NT pro BNP, and peptides which are derived from BNP, pre pro BNP and NT pro BNP and which are present in the blood as a result of proteolyses of BNP, NT pro BNP and pre pro BNP, are collectively described as markers related to or associated with BNP.

[0142] The term "BNP" as used herein refers to the mature 32-amino acid BNP molecule itself. As the skilled artisan will recognize, however, because of its relationship to BNP, the concentration of NT pro-BNP molecule can also provide diagnostic or prognostic information in patients. The phrase "marker related to BNP or BNP related peptide" refers to any polypeptide that originates from the pre pro-BNP molecule, other than the 32-amino acid BNP molecule itself. Proteolytic degradation of BNP and of peptides related to BNP have also been described in the literature and these proteolytic fragments are also encompassed by the term "BNP related peptides."

[0143] BNP and BNP-related peptides are predominantly found in the secretory granules of the cardiac ventricles, and are released from the heart in response to both ventricular volume expansion and pressure overload. Wilkins, M. *et al.*, *Lancet* 349: 1307-10 (1997). Elevations of BNP are associated with raised atrial and pulmonary wedge pressures, reduced ventricular systolic and diastolic function, left ventricular hypertrophy, and myocardial infarction. Sagnella, G.A., *Clinical Science* 95: 519-29 (1998). Furthermore, there are numerous reports of elevated BNP concentration associated with congestive heart failure and renal failure. Thus, BNP levels in a patient may be indicative of several possible underlying causes of dyspnea.

[0144] *D-dimer*

[0145] *D-dimer* is a crosslinked fibrin degradation product with an approximate molecular mass of 200 kDa. The normal plasma concentration of D-dimer is < 150 ng/ml

(750 pM). The plasma concentration of D-dimer is elevated in patients with acute myocardial infarction and unstable angina, but not stable angina. Hoffmeister, H.M. *et al.*, *Circulation* 91: 2520-27 (1995); Bayes-Genis, A. *et al.*, *Thromb. Haemost.* 81: 865-68 (1999); Gurfinkel, E. *et al.*, *Br. Heart J.* 71: 151-55 (1994); Kruskal, J.B. *et al.*, *N. Engl. J. Med.* 317: 1361-65 (1987); Tanaka, M. and Suzuki, A., *Thromb. Res.* 76: 289-98 (1994).

[0146] The plasma concentration of D-dimer also will be elevated during any condition associated with coagulation and fibrinolysis activation, including stroke, surgery, atherosclerosis, trauma, and thrombotic thrombocytopenic purpura. D-dimer is released into the bloodstream immediately following proteolytic clot dissolution by plasmin. The plasma concentration of D-dimer can exceed 2 µg/ml in patients with unstable angina. Gurfinkel, E. *et al.*, *Br. Heart J.* 71: 151-55 (1994). Plasma D-dimer is a specific marker of fibrinolysis and indicates the presence of a prothrombotic state associated with acute myocardial infarction and unstable angina. The plasma concentration of D-dimer is also nearly always elevated in patients with acute pulmonary embolism; thus, normal levels of D-dimer may allow the exclusion of pulmonary embolism. Egermayer *et al.*, *Thorax* 53: 830-34 (1998).

#### [0147] Cardiac Troponin

[0148] Troponin I (TnI) is a 25 kDa inhibitory element of the troponin complex, found in muscle tissue. TnI binds to actin in the absence of Ca<sup>2+</sup>, inhibiting the ATPase activity of actomyosin. A TnI isoform that is found in cardiac tissue (cTnI) is 40% divergent from skeletal muscle TnI, allowing both isoforms to be immunologically distinguished. The normal plasma concentration of cTnI is < 0.1 ng/ml (4 pM). cTnI is released into the bloodstream following cardiac cell death; thus, the plasma cTnI concentration is elevated in patients with acute myocardial infarction. Investigations into changes in the plasma cTnI concentration in patients with unstable angina have yielded mixed results, but cTnI is not elevated in the plasma of individuals with stable angina. Benamer, H. *et al.*, *Am. J. Cardiol.* 82: 845-50 (1998); Bertinchant, J.P. *et al.*, *Clin. Biochem.* 29: 587-94 (1996); Tanasijevic, M.J. *et al.*, *Clin. Cardiol.* 22: 13-16 (1999); Musso, P. *et al.*, *J. Ital. Cardiol.* 26: 1013-23 (1996); Holvoet, P. *et al.*, *JAMA* 281: 1718-21 (1999); Holvoet, P. *et al.*, *Circulation* 98: 1487-94 (1998).

[0149] The plasma concentration of cTnI in patients with acute myocardial infarction is significantly elevated 4-6 hours after onset, peaks between 12-16 hours, and can remain elevated for one week. The release kinetics of cTnI associated with unstable angina may be similar. The measurement of specific forms of cardiac troponin, including free cardiac troponin I and complexes of cardiac troponin I with troponin C and/or T may provide the user with the ability to identify various stages of ACS. Free and complexed cardiac troponin T may be used in a manner analogous to that described for cardiac troponin I. Cardiac troponin T complex may be useful either alone or when expressed as a ratio with total cardiac troponin I to provide information related to the presence of progressing myocardial damage. Ongoing ischemia may result in the release of the cardiac troponin TIC complex, indicating that higher ratios of cardiac troponin TIC:total cardiac troponin I may be indicative of continual damage caused by unresolved ischemia. *See*, U.S. Patent Nos. 6,147,688, 6,156,521, 5,947,124, and 5,795,725.

[0150] Based on the foregoing discussion, the skilled artisan will recognize that, for example, increased BNP is indicative of congestive heart failure, but may also be indicative of other cardiac-related conditions such as myocardial infarction. Thus, the inclusion of a marker related to myocardial injury such as cardiac troponin I and/or cardiac troponin T can permit further discrimination of the disease underlying the observed dyspnea and the increased BNP level. In this case, an increased level of cardiac troponin may be used to rule in myocardial infarction.

[0151] Similarly, BNP may also be indicative of pulmonary embolism. The inclusion of a marker related to coagulation and hemostasis such as D-dimer can permit further discrimination of the disease underlying the observed dyspnea and the increased BNP level. In this case, a normal level of D-dimer may be used to rule out pulmonary embolism.

[0152] The skilled artisan will readily acknowledge that other markers may be substituted in or added to this marker panel to further discriminate the causes of dyspnea. Suitable markers are described in co-pending PCT Application No. \_\_\_\_\_, filed December 23, 2003 (Atty. Docket No. 071949-5603), which is hereby incorporated by reference in its entirety. Preferred panels for the diagnosis of a cause of dyspnea comprise a plurality of markers independently selected from the group consisting of specific

markers of cardiac injury, specific markers of neural tissue injury, non-specific markers of tissue injury, markers related to blood pressure regulation, markers related to inflammation, markers related to coagulation and hemostasis, markers related to pulmonary injury, and markers related to apoptosis. Exemplary markers in each of these groups are described hereinafter. Preferably, such a panel comprises markerd from two, three, four, five, or more different members of this group. Thus, particularly preferred panels for the diagnosis of a cause of dyspnea comprise one or more specific markers of cardiac injury and one or more markers related to blood pressure regulation; one or more specific markers of cardiac injury and one or more markers related to coagulation and hemostasis; one or more markers related to blood pressure regulation and one or more markers related to coagulation and hemostasis; or one or more specific markers of cardiac injury, one or more markers related to blood pressure regulation, and one or more markers related to coagulation and hemostasis, where each of these particularly preferred panels may optionally comprise one or more non-specific markers of tissue injury, markers related to inflammation, markers related to pulmonary injury, and/or markers related to apoptosis.

[0153] In similar fashion, a panel may comprise a plurality of markers selected to diagnose, and/or distinguish amongst a plurality of, cerebrovascular disorders. In these aspects related to cerebrovascular disease, preferred marker panels comprise a plurality of markers independently selected from the group consisting of specific markers of neural tissue injury, markers related to blood pressure regulation, markers related to coagulation and hemostasis, markers related to inflammation, and markers related to apoptosis. Preferably, such a panel comprises markerd from two, three, four, or five different members of this group.

[0154] The following table provides an exemplary list of markers for use in the methods described herein:

<u>Marker</u>	<u>Classification</u>
Myoglobin	Nonspecific tissue injury
E-selectin	Nonspecific tissue injury
VEGF	Nonspecific tissue injury
Troponin I and complexes	Myocardial injury
Troponin T and complexes	Myocardial injury
Annexin V	Myocardial injury
B-enolase	Myocardial injury

CK-MB	Myocardial injury
Glycogen phosphorylase-BB	Myocardial injury
Heart type fatty acid binding protein	Myocardial injury
Phosphoglyceric acid mutase	Myocardial injury
S-100ao	Myocardial injury
ANP	Blood pressure regulation
CNP	Blood pressure regulation
urotensin II	Blood pressure regulation
BNP	Blood pressure regulation
calcitonin gene related peptide	Blood pressure regulation
arg-Vasopressin	Blood pressure regulation
Endothelin-1	Blood pressure regulation
Endothelin-2	Blood pressure regulation
Endothelin-31	Blood pressure regulation
procalcitonin	Blood pressure regulation
calcyphosine	Blood pressure regulation
adrenomedullin	Blood pressure regulation
aldosterone	Blood pressure regulation
angiotensin 1	Blood pressure regulation
angiotensin 2	Blood pressure regulation
angiotensin 3	Blood pressure regulation
Bradykinin	Blood pressure regulation
calcitonin	Blood pressure regulation
Endothelin-2	Blood pressure regulation
Endothelin-3	Blood pressure regulation
Renin	Blood pressure regulation
Urodilatin	Blood pressure regulation
Plasmin	Coagulation and hemostasis
Thrombin	Coagulation and hemostasis
Antithrombin-III	Coagulation and hemostasis
Fibrinogen	Coagulation and hemostasis
von Willebrand factor	Coagulation and hemostasis
D-dimer	Coagulation and hemostasis
PAI-1	Coagulation and hemostasis
PROTEIN C	Coagulation and hemostasis
TAFI	Coagulation and hemostasis
Fibrinopeptide A	Coagulation and hemostasis
Plasmin alpha 2 antiplasmin complex	Coagulation and hemostasis
Platelet factor 4	Coagulation and hemostasis
Platelet-derived growth factor	Coagulation and hemostasis
P-selectin	Coagulation and hemostasis
Prothrombin fragment 1+2	Coagulation and hemostasis
B-thromboglobulin	Coagulation and hemostasis
Thrombin antithrombin III complex	Coagulation and hemostasis
Thrombomodulin	Coagulation and hemostasis
Thrombus Precursor Protein	Coagulation and hemostasis
Tissue factor	Coagulation and hemostasis

basic calponin 1	Vascular tissue
beta like 1 integrin	Vascular tissue
Calponin	Vascular tissue
CSRP2	Vascular tissue
elastin	Vascular tissue
Fibrillin 1	Vascular tissue
LTBP4	Vascular tissue
smooth muscle myosin	Vascular tissue
transgelin	Vascular tissue
Carboxyterminal propeptide of type I procollagen (PICP)	Collagen synthesis
Collagen carboxyterminal telopeptide (ICTP)	Collagen degradation
Glutathione S Transferase	Inflammatory
HIF 1 ALPHA	Inflammatory
IL-10	Inflammatory
IL-1-Beta	Inflammatory
IL-1ra	Inflammatory
IL-6	Inflammatory
IL-8	Inflammatory
Lysophosphatidic acid	Inflammatory
MDA-modified LDL	Inflammatory
Human neutrophil elastase	Inflammatory
C-reactive protein	Inflammatory
Insulin-like growth factor	Inflammatory
Inducible nitric oxide synthase	Inflammatory
Intracellular adhesion molecule	Inflammatory
Lactate dehydrogenase	Inflammatory
MCP-1	Inflammatory
MDA-LDL	Inflammatory
MMP-1	Inflammatory
MMP-2	Inflammatory
MMP-3	Inflammatory
MMP-9	Inflammatory
TIMP-1	Inflammatory
TIMP-2	Inflammatory
TIMP-3	Inflammatory
n-acetyl aspartate	Inflammatory
TNF Receptor Superfamily Member 1A	Inflammatory
Transforming growth factor beta	Inflammatory
Tumor necrosis factor alpha	Inflammatory
Vascular cell adhesion molecule	Inflammatory
Vascular endothelial growth factor	Inflammatory
cystatin C	Inflammatory
substance P	Inflammatory
Myceloperoxidase (MPO)	Inflammatory
macrophage inhibitory factor	Inflammatory
Fibronectin	Inflammatory

cardiotrophin 1	Inflammatory
Haptoglobin	Inflammatory
PAPPA	Inflammatory
s-CD40 ligand*	Inflammatory
HMG	Inflammatory
IL -1	Inflammatory
IL -2	Inflammatory
IL -4	Inflammatory
IL -6	Inflammatory
IL-8	Inflammatory
IL -10	Inflammatory
IL -11	Inflammatory
IL -13	Inflammatory
IL -18	Inflammatory
Eosinophil cationic protein	Inflammatory
Mast cell tryptase	Inflammatory
VCAM	Inflammatory
sICAM-1	Inflammatory
TNF $\alpha$	Inflammatory
Osteoprotegerin	Inflammatory
Prostaglandin D-synthase	Inflammatory
Prostaglandin E2	Inflammatory
RANK ligand	Inflammatory
HSP-60	Inflammatory
Serum Amyloid A	Inflammatory
s-iL 18 receptor	Inflammatory
S-iL-1 receptor	Inflammatory
s-TNF P55	Inflammatory
s-TNF P75	Inflammatory
TGF-beta	Inflammatory
MMP-11	Inflammatory
Beta NGF	Inflammatory
CD44	Inflammatory
EGF	Inflammatory
E-selectin	Inflammatory
Fibronectin	Inflammatory
Neutrophil elastase	Pulmonary injury
KL-6	Pulmonary injury
LAMP 3	Pulmonary injury
LAMP3	Pulmonary injury
Lung Surfactant protein A	Pulmonary injury
Lung Surfactant protein B	Pulmonary injury
Lung Surfactant protein C	Pulmonary injury
Lung Surfactant protein D	Pulmonary injury
phospholipase D	Pulmonary injury
PLA2G5	Pulmonary injury
SFTPC	Pulmonary injury



MAPK10	Neural tissue injury
KCNK4	Neural tissue injury
KCNK9	Neural tissue injury
KCNQ5	Neural tissue injury
14-3-3	Neural tissue injury
4.1B	Neural tissue injury
APO E4-1	Neural tissue injury
myelin basic protein	Neural tissue injury
Atrophin 1	Neural tissue injury
brain Derived neurotrophic factor	Neural tissue injury
Brain Fatty acid binding protein	Neural tissue injury
brain tubulin	Neural tissue injury
CACNA1A	Neural tissue injury
Calbindin D	Neural tissue injury
Calbrain	Neural tissue injury
Carbonic anhydrase XI	Neural tissue injury
CBLN1	Neural tissue injury
Cerebellin 1	Neural tissue injury
Chimerin 1	Neural tissue injury
Chimerin 2	Neural tissue injury
CHN1	Neural tissue injury
CHN2	Neural tissue injury
Ciliary neurotrophic factor	Neural tissue injury
CK-BB	Neural tissue injury
CRHR1	Neural tissue injury
C-tau	Neural tissue injury
DRPLA	Neural tissue injury
GFAP	Neural tissue injury
GPM6B	Neural tissue injury
GPR7	Neural tissue injury
GPR8	Neural tissue injury
GRIN2C	Neural tissue injury
GRM7	Neural tissue injury
HAPIP	Neural tissue injury
HIP2	Neural tissue injury
LDH	Neural tissue injury
Myelin basic protein	Neural tissue injury
NCAM	Neural tissue injury
NT-3	Neural tissue injury
NDPKA	Neural tissue injury
Neural cell adhesion molecule	Neural tissue injury
NEUROD2	Neural tissue injury
Neurofilament L	Neural tissue injury
Neuroglobin	Neural tissue injury
neuromodulin	Neural tissue injury
Neuron specific enolase	Neural tissue injury
Neuropeptide Y	Neural tissue injury

Neurotensin	Neural tissue injury
Neurotrophin 1,2,3,4	Neural tissue injury
NRG2	Neural tissue injury
PACE4	Neural tissue injury
phosphoglycerate mutase	Neural tissue injury
PKC gamma	Neural tissue injury
proteolipid protein	Neural tissue injury
PTEN	Neural tissue injury
PTPRZ1	Neural tissue injury
RGS9	Neural tissue injury
RNA Binding protein Regulatory Subunit	Neural tissue injury
S-100 $\beta$	Neural tissue injury
SCA7	Neural tissue injury
secretagogin	Neural tissue injury
SLC1A3	Neural tissue injury
SORL1	Neural tissue injury
SREB3	Neural tissue injury
STAC	Neural tissue injury
STX1A	Neural tissue injury
STXBP1	Neural tissue injury
Syntaxin	Neural tissue injury
thrombomodulin	Neural tissue injury
transferrin	Neural tissue injury
adenylate kinase-1	Neural tissue injury
BDNF*	Neural tissue injury
neurokinin A	Neural tissue injury
s-acetyl Glutathione	apoptosis
cytochrome C	apoptosis
Caspase 3	apoptosis
Cathepsin D	apoptosis
$\alpha$ -spectrin	apoptosis

[0155] Assay Measurement Strategies

[0156] Numerous methods and devices are well known to the skilled artisan for the detection and analysis of the markers of the instant invention. With regard to polypeptides or proteins in patient test samples, immunoassay devices and methods are often used. See, e.g., U.S. Patents 6,143,576; 6,113,855; 6,019,944; 5,985,579; 5,947,124; 5,939,272; 5,922,615; 5,885,527; 5,851,776; 5,824,799; 5,679,526; 5,525,524; and 5,480,792, each of which is hereby incorporated by reference in its entirety, including all tables, figures and claims. These devices and methods can utilize labeled molecules in various sandwich, competitive, or non-competitive assay formats, to generate a signal that is related to the

presence or amount of an analyte of interest. Additionally, certain methods and devices, such as biosensors and optical immunoassays, may be employed to determine the presence or amount of analytes without the need for a labeled molecule. *See, e.g.*, U.S. Patents 5,631,171; and 5,955,377, each of which is hereby incorporated by reference in its entirety, including all tables, figures and claims. One skilled in the art also recognizes that robotic instrumentation including but not limited to Beckman Access, Abbott AxSym, Roche ElecSys, Dade Behring Stratus systems are among the immunoassay analyzers that are capable of performing the immunoassays taught herein.

[0157] Preferably the markers are analyzed using an immunoassay, although other methods are well known to those skilled in the art (for example, the measurement of marker RNA levels). The presence or amount of a marker is generally determined using antibodies specific for each marker and detecting specific binding. Any suitable immunoassay may be utilized, for example, enzyme-linked immunoassays (ELISA), radioimmunoassays (RIAs), competitive binding assays, and the like. Specific immunological binding of the antibody to the marker can be detected directly or indirectly. Direct labels include fluorescent or luminescent tags, metals, dyes, radionuclides, and the like, attached to the antibody. Indirect labels include various enzymes well known in the art, such as alkaline phosphatase, horseradish peroxidase and the like.

[0158] The use of immobilized antibodies specific for the markers is also contemplated by the present invention. The antibodies could be immobilized onto a variety of solid supports, such as magnetic or chromatographic matrix particles, the surface of an assay plate (such as microtiter wells), pieces of a solid substrate material or membrane (such as plastic, nylon, paper), and the like. An assay strip could be prepared by coating the antibody or a plurality of antibodies in an array on solid support. This strip could then be dipped into the test sample and then processed quickly through washes and detection steps to generate a measurable signal, such as a colored spot.

[0159] The analysis of a plurality of markers may be carried out separately or simultaneously with one test sample. For separate or sequential assay of markers, suitable apparatuses include clinical laboratory analyzers such as the ElecSys (Roche), the AxSym (Abbott), the Access (Beckman), the ADVIA® CENTAUR® (Bayer) immunoassay systems, the NICHOLS ADVANTAGE® (Nichols Institute) immunoassay system, etc.

Preferred apparatuses or protein chips perform simultaneous assays of a plurality of markers on a single surface. Particularly useful physical formats comprise surfaces having a plurality of discrete, addressable locations for the detection of a plurality of different analytes. Such formats include protein microarrays, or "protein chips" (*see, e.g., Ng and Ilag, J. Cell Mol. Med.* 6:329-340 (2002)) and certain capillary devices (*see, e.g., U.S. Patent No. 6,019,944*). In these embodiments, each discrete surface location may comprise antibodies to immobilize one or more analyte(s) (*e.g., a marker*) for detection at each location. Surfaces may alternatively comprise one or more discrete particles (*e.g., microparticles or nanoparticles*) immobilized at discrete locations of a surface, where the microparticles comprise antibodies to immobilize one analyte (*e.g., a marker*) for detection.

[0160] Several markers may be combined into one test for efficient processing of a multiple of samples. In addition, one skilled in the art would recognize the value of testing multiple samples (for example, at successive time points) from the same individual. Such testing of serial samples will allow the identification of changes in marker levels over time. Increases or decreases in marker levels, as well as the absence of change in marker levels, would provide useful information about the disease status that includes, but is not limited to identifying the approximate time from onset of the event, the presence and amount of salvagable tissue, the appropriateness of drug therapies, the effectiveness of various therapies as indicated by reperfusion or resolution of symptoms, differentiation of the various types of ACS, identification of the severity of the event, identification of the disease severity, and identification of the patient's outcome, including risk of future events.

[0161] A panel consisting of the markers referenced above may be constructed to provide relevant information related to differential diagnosis and/or prognosis. Such a panel may be constructed using 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or more or individual markers. The analysis of a single marker or subsets of markers comprising a larger panel of markers could be carried out by one skilled in the art to optimize clinical sensitivity or specificity in various clinical settings. These include, but are not limited to ambulatory, urgent care, critical care, intensive care, monitoring unit, inpatient, outpatient, physician office, medical clinic, and health screening settings. Furthermore, one skilled in the art can use a single marker or a subset of markers comprising a larger panel of markers in

combination with an adjustment of the diagnostic threshold in each of the aforementioned settings to optimize clinical sensitivity and specificity. The clinical sensitivity of an assay is defined as the percentage of those with the disease that the assay correctly predicts, and the specificity of an assay is defined as the percentage of those without the disease that the assay correctly predicts (Tietz Textbook of Clinical Chemistry, 2<sup>nd</sup> edition, Carl Burtis and Edward Ashwood eds., W.B. Saunders and Company, p. 496).

[0162] The analysis of markers could be carried out in a variety of physical formats as well. For example, the use of microtiter plates or automation could be used to facilitate the processing of large numbers of test samples. Alternatively, single sample formats could be developed to facilitate immediate treatment and diagnosis in a timely fashion, for example, in ambulatory transport or emergency room settings.

[0163] In another embodiment, the present invention provides a kit for the analysis of markers. Such a kit preferably comprises devices and reagents for the analysis of at least one test sample and instructions for performing the assay. Optionally the kits may contain one or more means for using information obtained from immunoassays performed for a marker panel to rule in or out certain diagnoses.

[0164] Selection of Antibodies

[0165] The generation and selection of antibodies may be accomplished several ways. For example, one way is to purify polypeptides of interest or to synthesize the polypeptides of interest using, e.g., solid phase peptide synthesis methods well known in the art. See, e.g., *Guide to Protein Purification*, Murray P. Deutcher, ed., *Meth. Enzymol.* Vol 182 (1990); *Solid Phase Peptide Synthesis*, Greg B. Fields ed., *Meth. Enzymol.* Vol 289 (1997); Kiso *et al.*, *Chem. Pharm. Bull. (Tokyo)* 38: 1192-99, 1990; Mostafavi *et al.*, *Biomed. Pept. Proteins Nucleic Acids* 1: 255-60, 1995; Fujiwara *et al.*, *Chem. Pharm. Bull. (Tokyo)* 44: 1326-31, 1996. The selected polypeptides may then be injected, for example, into mice or rabbits, to generate polyclonal or monoclonal antibodies. One skilled in the art will recognize that many procedures are available for the production of antibodies, for example, as described in *Antibodies, A Laboratory Manual*, Ed Harlow and David Lane, Cold Spring Harbor Laboratory (1988), Cold Spring Harbor, N.Y. One skilled in the art will also appreciate that binding fragments or Fab fragments which mimic antibodies can also be prepared from genetic information by various procedures (Antibody

Engineering: A Practical Approach (Borrebaeck, C., ed.), 1995, Oxford University Press, Oxford; J. Immunol. 149, 3914-3920 (1992)).

[0166] In addition, numerous publications have reported the use of phage display technology to produce and screen libraries of polypeptides for binding to a selected target. See, e.g., Cwirla *et al.*, *Proc. Natl. Acad. Sci. USA* 87, 6378-82, 1990; Devlin *et al.*, *Science* 249, 404-6, 1990; Scott and Smith, *Science* 249, 386-88, 1990; and Ladner *et al.*, U.S. Pat. No. 5,571,698. A basic concept of phage display methods is the establishment of a physical association between DNA encoding a polypeptide to be screened and the polypeptide. This physical association is provided by the phage particle, which displays a polypeptide as part of a capsid enclosing the phage genome which encodes the polypeptide. The establishment of a physical association between polypeptides and their genetic material allows simultaneous mass screening of very large numbers of phage bearing different polypeptides. Phage displaying a polypeptide with affinity to a target bind to the target and these phage are enriched by affinity screening to the target. The identity of polypeptides displayed from these phage can be determined from their respective genomes. Using these methods a polypeptide identified as having a binding affinity for a desired target can then be synthesized in bulk by conventional means. See, e.g., U.S. Patent No. 6,057,098, which is hereby incorporated in its entirety, including all tables, figures, and claims.

[0167] The antibodies that are generated by these methods may then be selected by first screening for affinity and specificity with the purified polypeptide of interest and, if required, comparing the results to the affinity and specificity of the antibodies with polypeptides that are desired to be excluded from binding. The screening procedure can involve immobilization of the purified polypeptides in separate wells of microtiter plates. The solution containing a potential antibody or groups of antibodies is then placed into the respective microtiter wells and incubated for about 30 min to 2 h. The microtiter wells are then washed and a labeled secondary antibody (for example, an anti-mouse antibody conjugated to alkaline phosphatase if the raised antibodies are mouse antibodies) is added to the wells and incubated for about 30 min and then washed. Substrate is added to the wells and a color reaction will appear where antibody to the immobilized polypeptide(s) are present.

[0168] The antibodies so identified may then be further analyzed for affinity and specificity in the assay design selected. In the development of immunoassays for a target protein, the purified target protein acts as a standard with which to judge the sensitivity and specificity of the immunoassay using the antibodies that have been selected. Because the binding affinity of various antibodies may differ, certain antibody pairs (*e.g.*, in sandwich assays) may interfere with one another sterically, *etc.*, assay performance of an antibody may be a more important measure than absolute affinity and specificity of an antibody.

[0169] Those skilled in the art will recognize that many approaches can be taken in producing antibodies or binding fragments and screening and selecting for affinity and specificity for the various polypeptides, but these approaches do not change the scope of the invention.

[0170] Selecting a Treatment Regimen

[0171] The appropriate treatments for various types of vascular disease may be large and diverse. However, once a diagnosis is obtained, the clinician can readily select a treatment regimen that is compatible with the diagnosis. Accordingly, the present invention provides methods of early differential diagnosis to allow for appropriate intervention in acute time windows. The skilled artisan is aware of appropriate treatments for numerous diseases discussed in relation to the methods of diagnosis described herein. *See, e.g., Merck Manual of Diagnosis and Therapy*, 17<sup>th</sup> Ed. Merck Research Laboratories, Whitehouse Station, NJ, 1999.

[0172] The following provides a brief discussion of additional exemplary markers for use in identifying suitable marker panels by the methods described herein.

[0173] Examples:

[0174] Example 1: Selection of Markers for a Stroke Panel. A set of samples from patients diagnosed with stroke and normal healthy donors were assayed for several markers of potential utility. No individual marker has sufficient clinical utility to diagnose stroke. The methods described above were used to determine the optimum markers for use in a panel of markers. The data was separated into diseased and non-diseased groups. The indicator functions were selected to be ramp functions for all markers. The objective

function was chosen to be the product of the area, the knee, the specificity at 92.5% sensitivity and the sensitivity at 92.5% specificity. The initial simplex was randomly distributed about a vertex derived from the univariate analysis. Using the downhill simplex method with simulated annealing a local minimum was found that maximized the objective function. For contribution for each analyte was calculated by setting the weighting parameter to zero and calculating the change in the objective function. This process was repeated 50 times. The markers were ranked by their average contribution over the 50 optimizations. The ROC curves for the initial vertex and an optimization are shown in figure 10. The ranking of the marker contributions is shown in figure 11. The lowest half of the markers were removed from the panel and the process was repeated. Figure 12 and 13 show the same information as in figures 10 and 11 but for the 19 marker panel. The lowest 9 markers were removed from the panel and the process was repeated. Figures 14 and 15 show the same information as in Figures 10 and 11 but for the 10 marker panel. The lowest 5 markers were removed from the panel and the process was repeated a final time. Figure 16 and 17 show the same information as in figures 10 and 11 but for the 5 marker panel. The individual ROC curves of the final 5 markers are shown in figure 9. The order of the contribution does not match the order of the area of the individual ROC curves. A marker with poorer univariate utility may have greater utility when used in a panel. The area of the ROC curve decreases with decreasing panel size.

[0175] Example 2: Improvement in diagnosis of AMI Utilizing Changes in Marker Levels. Data from a clinical study from patients presenting with chest pain with serial draws from each patient was analyzed using the methods described in this document. The data was first analyzed without using derived markers. The data was again analyzed utilizing derived markers that were related to the change in marker value from the initial value. The ROC curves from both optimized panel responses are shown in figure 18. The data clearly illustrates the utility of the change in markers to improve the diagnostic ability of panels in acute disease states. The method was also applied to determine the best 3 and 2 marker panels, and the results are also shown in figure 18. Figure 19 shows the contributions of the six AMI markers. Myoglobin, while not a specific marker for AMI is a small molecule and the first marker of the three to elevate after AMI. TnI is a specific marker for AMI, but is released more slowly. The method was not aware of this but still chose TnI value and change in Myoglobin.



[0176] Example 3: Simultaneous Optimization of Two Criteria. In this example known stroke samples are analyzed with both stroke mimics and NHD samples in the non-diseased set. There are about 50 mimics and about 500 NHD samples, so the weighting is heavily in favor of optimizing results for NHD samples. After optimization the panel response is applied to a test set stroke vs. mimics and stroke vs. NHD. Similarly the data was optimized on stroke vs mimics, and the panel response was applied to as test set stroke vs. NHD and stroke vs NHD and mimics. Table 2 shows the average results of sample runs applied to the optimization sets and then to the test sets. The effectiveness of the test is poor with respect to mimics. Two more optimizations were made as before, but this time a second group of data is simultaneously optimized. The second group consists of the stroke samples and the mimics. Table 2 also shows the average results of sample runs applied to the optimization set and when the panel response is applied to the two test sets. The effectiveness of the test with respect to mimics is now improved.

[0177] While the invention has been described and exemplified in sufficient detail for those skilled in this art to make and use it, various alternatives, modifications, and improvements should be apparent without departing from the spirit and scope of the invention.

[0178] One skilled in the art readily appreciates that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The examples provided herein are representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention. Modifications therein and other uses will occur to those skilled in the art. These modifications are encompassed within the spirit of the invention and are defined by the scope of the claims.

[0179] It will be readily apparent to a person skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention.

[0180] All patents and publications mentioned in the specification are indicative of the levels of those of ordinary skill in the art to which the invention pertains. All patents and publications are herein incorporated by reference to the same extent as if each individual publication was specifically and individually indicated to be incorporated by reference.

[0181] The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations which is not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising", "consisting essentially of" and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the appended claims.

[0182] Other embodiments are set forth within the following claims.

TABLE 1

MARKER	SENSE OF MARKER	CUTOFF LOCATION	LENGTH OF CUTOFF	WEIGHTING COEFF.
Analyte 1	Incr.	18.01	0.90	0.67
Analyte 2	Incr.	128.92	0.83	0.75
Analyte 3	Incr.	86.17	1.00	0.73
Analyte 4	Incr.	41.46	0.99	0.55
Analyte 5	Incr.	238.23	1.00	0.74
Analyte 6	Incr.	21.87	1.00	0.82
Analyte 7	Incr.	2.63	0.96	0.66
Analyte 8	Decr.	65.92	0.14	0.66
Analyte 9	Incr.	582.80	0.82	0.57
Analyte 10	Incr.	66.07	1.00	0.65
Analyte 11	Incr.	0.00	1.00	0.81
Analyte 12	Incr.	189.17	0.57	0.84
Analyte 13	Incr.	122.76	1.00	0.68
Analyte 14	Incr.	45.72	1.00	0.67
Analyte 15	Incr.	1632.97	1.00	0.72
Analyte 16	Incr.	48.93	0.74	0.82
Analyte 17	Incr.	8352.03	0.18	0.85
Analyte 18	Incr.	4528.32	0.18	0.78
Analyte 19	Incr.	1424.02	0.56	0.83
Analyte 20	Incr.	1827.05	0.49	0.84
Analyte 21	Incr.	5856.94	0.68	0.73
Analyte 22	Incr.	58.83	1.00	0.57
Analyte 23	Incr.	4556.97	0.71	0.63
Analyte 24	Incr.	224.83	0.41	0.68
Analyte 25	Incr.	10080.59	0.89	0.53
Analyte 26	Incr.	13.74	0.50	0.66
Analyte 27	Incr.	2.64	0.43	0.77
Analyte 28	Decr.	11678.95	0.69	0.52
Analyte 29	Incr.	1.70	1.00	0.66
Analyte 30	Incr.	1283.89	1.00	0.54
Analyte 31	Incr.	10.96	1.00	0.50
Analyte 32	Decr.	18882.79	1.00	0.58
Analyte 33	Decr.	0.42	1.00	0.62
Analyte 34	Decr.	3.99	0.96	0.52
Analyte 35	Decr.	4950.62	0.41	0.64
Analyte 36	Incr.	45.17	1.00	0.52
Analyte 37	Incr.	126.85	0.71	0.58
Analyte 38	Decr.	686.75	0.47	0.73

We claim:

1. A method of identifying a panel of markers for diagnosis or prognosis of a disease or a condition, comprising:

a) calculating a panel response for each patient in a set of diseased patients and in a set of non-diseased patients, said panel response being a function of value of each of a plurality of markers in a panel of markers;

b) calculating a value for an objective function, said objective function being indicative of an effectiveness of the panel; and

c) iterating steps a) and b) by varying at least one of parameters relating to said panel response function and a sense of each marker to facilitate optimization of said objective function.

2. The method according to claim 1, wherein said objective function is a measure of an overlap of panel responses of diseased patients and panel responses of non-diseased patients.

3. The method according to claim 1, wherein said panel response is a function of value of an indicator for each of a plurality of markers in a panel of markers and a weighting coefficient for each marker, said indicator being a mapping, for each of said plurality of markers, of marker levels, said mapping being according to an indicator function; and

wherein said iterating includes varying at least one of said weighting coefficients, parameters relating to said indicator function, and a sense of each marker to facilitate optimization of said objective function.

4. The method according to claim 3, wherein each indicator has a first value for marker levels below a cutoff region and a second value for marker values above a cutoff region, said cutoff region being defined by a location and a length.

5. The method according to claim 4, wherein said parameters include said location of said cutoff region and said length of said cutoff region.

6. The method according to claim 4, wherein said length of said cutoff region is zero.
7. The method according to claim 4, wherein said length of said cutoff region is greater than zero.
8. The method according to claim 7, wherein said indicators have values between said first value and said second value for marker levels within said cutoff region.
9. The method according to claim 8, wherein said indicators have values varying linearly from said first value to said second value across said cutoff region.
10. The method according to claim 8, wherein said indicators have values varying non-linearly from said first value to said second value across said cutoff region.
11. The method according to claim 10, wherein said non-linear variation is indicative of an error function of a distribution of marker values of diseased patients and an error function of a distribution of marker values of non-diseased patients within said cutoff region.
12. The method according to claim 3, wherein said calculating a panel response includes calculating, for each patient,  $\sum w_i I_i$ , where  $w$  is a weighting coefficient for a marker  $i$ ,  $I$  is the indicator value for the marker  $i$ , and  $\sum$  is a summation over all of said plurality of markers.
13. The method according to claim 1, wherein said calculating a value for an objective function includes generating a receiver operating characteristic (ROC) curve for said panel response, said ROC curve being indicative of a sensitivity of said panel response as a function of one minus a specificity of said panel response.
14. The method according to claim 13, wherein said objective function is associated with an area under said ROC curve.

15. The method according to claim 13, wherein said objective function is associated with a knee of said ROC curve.
16. The method according to claim 13, wherein said objective function is associated with a sensitivity at a selected specificity level.
17. The method according to claim 13, wherein said objective function is associated with a specificity at a selected sensitivity level.
18. The method according to claim 13, wherein said objective function is associated with two or more of an area under said ROC curve, a knee of said ROC curve, a sensitivity at a selected specificity level, and a specificity at a selected sensitivity level.
19. The method according to claim 13, wherein said iterating constrains at least one of an area under said ROC curve, a knee of said ROC curve, a sensitivity at a selected specificity level, and a specificity at a selected sensitivity level to be above about 0.9.
20. The method according to claim 1, further comprising:
  - d) removing at least one of said markers from said panel;
  - e) calculating a value of said objective function; and
  - f) determining a contribution of said at least one of said markers to said objective function based on a result of step e).
21. The method according to claim 20, further comprising:
  - g) repeating steps d) through f) by removing a different at least one of said markers from said panel; and
  - h) eliminating a marker from said panel of markers in accordance with said contribution of said marker to said objective function.
22. The method according to claim 1, further comprising:
  - d) removing at least one of said markers from said panel;
  - e) iterating steps a) and b) by varying parameters relating to said panel response function to facilitate optimization of said objective function; and

f) determining a contribution of said at least one of said markers to said objective function based on a result of step e).

23. The method according to claim 22, further comprising:

g) repeating steps d) through f) by removing a different at least one of said markers from said panel; and

h) eliminating a marker from said panel of markers in accordance with said contribution of said marker to said objective function.

24. A system for identifying a panel of markers for diagnosis or prognosis of a disease or a condition, comprising:

means for calculating a panel response for each patient in a set of diseased patients and in a set of non-diseased patients, said panel response being a function of value of each of a plurality of markers in a panel of markers;

means for calculating a value for an objective function, said objective function being indicative of an effectiveness of said panel; and

means for iteratively activating said means for calculating a panel response and said means for calculating a value for an objective function, by varying at least one of parameters relating to said panel response function and a sense of each marker to facilitate optimization of said objective function.

25. The system according to claim 24, wherein said objective function is a measure of an overlap of panel responses of diseased patients and panel responses of non-diseased patients.

26. The system according to claim 24, wherein said panel response is a function of value of an indicator for each of a plurality of markers in a panel of markers and a weighting coefficient for each marker, said indicator being a mapping, for each of said plurality of markers, of marker levels, said mapping being according to an indicator function; and

wherein said means for iteratively activating is adapted to vary at least one of said weighting coefficients, parameters relating to said indicator function, and a sense of each marker to facilitate optimization of said objective function.

27. The system according to claim 26, wherein each indicator has a first value for marker levels below a cutoff region and a second value for marker values above a cutoff region, said cutoff region being defined by a location and a length.
28. The method according to claim 27, wherein said parameters include said location of said cutoff region and said length of said cutoff region.
29. The system according to claim 27, wherein said length of said cutoff region is zero.
30. The system according to claim 27, wherein said length of said cutoff region is greater than zero.
31. The system according to claim 30, wherein said indicators have values between said first value and said second value for marker levels within said cutoff region.
32. The system according to claim 31, wherein said indicators have values varying linearly from said first value to said second value across said cutoff region.
33. The system according to claim 32, wherein said indicators have values varying non-linearly from said first value to said second value across said cutoff region.
34. The system according to claim 33, wherein said non-linear variation is indicative of an error function of a distribution of marker values of diseased patients and an error function of a distribution of marker values of non-diseased patients within said cutoff region.
35. The system according to claim 26, wherein said means for calculating a panel response is adapted to calculate, for each patient,  $\sum w_i I_i$ , where  $w$  is a weighting coefficient for a marker  $i$ ,  $I$  is the indicator value for the marker  $I$ , and  $\sum$  is a summation over all of said plurality of markers.



36. The system according to claim 24, wherein said means for calculating a value for an objective function is adapted to generate a receiver operating characteristic (ROC) curve for said panel response, said ROC curve being indicative of a sensitivity of said panel response as a function of one minus a specificity of said panel response.
37. The system according to claim 36, wherein said objective function is associated with an area under said ROC curve.
38. The system according to claim 36, wherein said objective function is associated with a knee of said ROC curve.
39. The system according to claim 38, wherein said objective function is associated with a sensitivity at a selected specificity level.
40. The system according to claim 36, wherein said objective function is associated with a specificity at a selected sensitivity level.
41. The system according to claim 36, wherein said objective function is associated with two or more of an area under said ROC curve, a knee of said ROC curve, a sensitivity at a selected specificity level, and a specificity at a selected sensitivity level.
42. The system according to claim 36, wherein said means for iteratively activating is adapted to constrain at least one of an area under said ROC curve, a knee of said ROC curve, a sensitivity at a selected specificity level, and a specificity at a selected sensitivity level to be above about 0.9.
43. The system according to claim 24, further comprising:  
means for determining a contribution of said at least one of said markers to said objective function, said means for determining being adapted to remove at least one of said markers from said panel and to activate said means for calculating a value for an objective function.

44. The system according to claim 43, further comprising:

means for eliminating a marker from said panel of markers in accordance with said contribution of said marker to said objective function, said means for eliminating being adapted to activate said means for determining a contribution by removing a different at least one of said markers from said panel.

45. The system according to claim 24, further comprising:

means for determining a contribution of said at least one of said markers to said objective function, said means for determining being adapted to remove at least one of said markers from said panel and to iteratively activate said means for calculating a panel response and said means for calculating a value for an objective function, by varying parameters relating to said panel response function to facilitate optimization of said objective function.

46. The system according to claim 45, further comprising:

means for eliminating a marker from said panel of markers in accordance with said contribution of said marker to said objective function, said means for eliminating being adapted to activate said means for determining a contribution by removing a different at least one of said markers from said panel.

47. A program product, comprising machine readable program code for causing a machine to perform following method steps:

- a) calculating a panel response for each patient in a set of diseased patients and in a set of non-diseased patients, said panel response being a function of value of each of a plurality of markers in a panel of markers;
- b) calculating a value for an objective function, said objective function being indicative of an effectiveness of said panel; and
- c) iterating steps a) and b) by varying at least one of parameters relating to said panel response function and a sense of each marker to facilitate optimization of said objective function.

48. The program product according to claim 47, wherein said objective function is a measure of an overlap of panel responses of diseased patients and panel responses of non-diseased patients.

49. The program product according to claim 47, wherein said panel response is a function of value of an indicator for each of a plurality of markers in a panel of markers and a weighting coefficient for each marker, said indicator being a mapping, for each of said plurality of markers, of marker levels, said mapping being according to an indicator function; and

wherein said iterating includes varying at least one of said weighting coefficients, parameters relating to said indicator function, and a sense of each marker to facilitate optimization of said objective function.

50. The program product according to claim 49, wherein each indicator has a first value for marker levels below a cutoff region and a second value for marker values above a cutoff region, said cutoff region being defined by a location and a length.

51. The program product according to claim 50, wherein said parameters include said location of said cutoff region and said length of said cutoff region.

52. The program product according to claim 50, wherein said length of said cutoff region is zero.

53. The program product according to claim 50, wherein said length of said cutoff region is greater than zero.

54. The program product according to claim 53, wherein said indicators have values between said first value and said second value for marker levels within said cutoff region.

55. The program product according to claim 54, wherein said indicators have values varying linearly from said first value to said second value across said cutoff region.

56. The program product according to claim 54, wherein said indicators have values varying non-linearly from said first value to said second value across said cutoff region.
57. The program product according to claim 56, wherein said non-linear variation is indicative of an error function of a distribution of marker values of diseased patients and an error function of a distribution of marker values of non-diseased patients within said cutoff region.
58. The program product according to claim 49, wherein said calculating a panel response includes calculating, for each patient,  $\sum w_i I_i$ , where  $w$  is a weighting coefficient for a marker  $i$ ,  $I$  is the indicator value for the marker  $i$ , and  $\sum$  is a summation over all of said plurality of markers.
59. The program product according to claim 47, wherein said calculating a value for an objective function includes generating a receiver operating characteristic (ROC) curve for said panel response, said ROC curve being indicative of a sensitivity of said panel response as a function of one minus a specificity of said panel response.
60. The program product according to claim 59, wherein said objective function is associated with an area under said ROC curve.
61. The program product according to claim 59, wherein said objective function is associated with a knee of said ROC curve.
62. The program product according to claim 59, wherein said objective function is associated with a sensitivity at a selected specificity level.
63. The program product according to claim 59, wherein said objective function is associated with a specificity at a selected sensitivity level.
64. The program product according to claim 59, wherein said objective function is associated with two or more of an area under said ROC curve, a knee of said ROC curve, a sensitivity at a selected specificity level, and a specificity at a selected sensitivity level.

65. The program product according to claim 59, wherein said iterating constrains at least one of an area under said ROC curve, a knee of said ROC curve, a sensitivity at a selected specificity level, and a specificity at a selected sensitivity level to be above about 0.9.

66. The program product according to claim 47, further comprising machine readable program code for causing a machine to perform following method steps:

- d) removing at least one of said markers from said panel;
- e) calculating a value of said objective function; and
- f) determining a contribution of said at least one of said markers to said

objective function based on a result of step e).

67. The program product according to claim 66, further comprising machine readable program code for causing a machine to perform following method steps:

- g) repeating steps d) through f) by removing a different at least one of said markers from said panel; and
- h) eliminating a marker from said panel of markers in accordance with said contribution of said marker to said objective function.

68. The program product according to claim 47, further comprising machine readable program code for causing a machine to perform following method steps:

- d) removing at least one of said markers from said panel;
- e) iterating steps a) and b) by varying parameters relating to said panel response function to facilitate optimization of said objective function; and
- f) determining a contribution of said at least one of said markers to said objective function based on a result of step e).

69. The program product according to claim 68, further comprising machine readable program code for causing a machine to perform following method steps:

- g) repeating steps d) through f) by removing a different at least one of said markers from said panel; and
- h) eliminating a marker from said panel of markers in accordance with said contribution of said marker to said objective function.

70. The program product according to claim 47, wherein said machine readable code is embedded in a portable meter.
71. The program product according to claim 70, wherein said portable meter is a fluorometer.
72. The program product according to claim 70, wherein said portable meter is a reflectometer.
73. The program product according to claim 47, wherein said machine readable code is embedded in a computer.
74. The program product according to claim 73, wherein said computer is a portable computer.
75. The program product according to claim 73, wherein said computer is adapted to be accessed through a network.
76. The program product according to claim 75, wherein said network is the Internet.
77. The program product according to claim 73, wherein said computer is adapted to be coupled to an analyzer.
78. The program product according to claim 77, wherein said analyzer is an immunoassay analyzer.
79. The program product according to claim 77, wherein said analyzer is a single nucleotide polymorphism detector.
80. The program product according to claim 77, wherein said analyzer is adapted to sort and count similar and different particles and cells.

81. A method of identifying a panel of markers for diagnosis or prognosis of a disease or a condition, comprising:

- a) selecting a panel of markers, said panel including a plurality of markers measured in a set of diseased patients and a set of non-diseased patients;
- b) defining a cutoff region of marker levels for each of said plurality of markers, said cutoff region having a location and a length;
- c) selecting a weighting coefficient for each of said plurality of markers;
- d) mapping, for each of said plurality of markers, marker levels to an indicator, each of said indicators having a first value for marker levels below said cutoff region and a second value for marker levels above said cutoff region;
- e) calculating a panel response for each patient in said set of diseased patients and in said set of non-diseased patients, said panel response being a function of value of said indicator for each marker and said weighting coefficient for each marker;
- f) calculating a value for an objective function, said objective function being indicative of an effectiveness of said panel; and
- g) iterating steps e) and f) by varying at least one of said location of said cutoff region, said length of said cutoff region, said weighting coefficients, and a sense of each marker to facilitate optimization of said objective function.

82. The method according to claim 81, wherein said objective function is a measure of an overlap of panel responses of diseased patients and panel responses of non-diseased patients.

83. The method according to claim 81, wherein said length of said cutoff region is zero.

84. The method according to claim 81, wherein said length of said cutoff region is greater than zero.

85. The method according to claim 84, wherein said indicators have values between said first value and said second value for marker levels within said cutoff region.

86. The method according to claim 85, wherein said indicators have values varying linearly from said first value to said second value across said cutoff region.
87. The method according to claim 85, wherein said indicators have values varying non-linearly from said first value to said second value across said cutoff region.
88. The method according to claim 87, wherein said non-linear variation is indicative of an error function of a distribution of marker values of diseased patients and an error function of a distribution of marker values of non-diseased patients within said cutoff region.
89. The method according to claim 81, wherein said calculating a panel response includes calculating, for each patient,  $\sum w_i I_i$ , where  $w$  is a weighting coefficient for a marker  $i$ ,  $I$  is the indicator value for the marker  $I$ , and  $\sum$  is a summation over all of said plurality of markers.
90. The method according to claim 81, wherein said calculating a value for an objective function includes generating a receiver operating characteristic (ROC) curve for said panel response, said ROC curve being indicative of a sensitivity of said panel response as a function of one minus a specificity of said panel response.
91. The method according to claim 90, wherein said objective function is associated with an area under said ROC curve.
92. The method according to claim 90, wherein said objective function is associated with a knee of said ROC curve.
93. The method according to claim 90, wherein said objective function is associated with a sensitivity at a selected specificity level.
94. The method according to claim 90, wherein said objective function is associated with a specificity at a selected sensitivity level.



95. The method according to claim 90, wherein said objective function is associated with two or more of an area under said ROC curve, a knee of said ROC curve, a sensitivity at a selected specificity level, and a specificity at a selected sensitivity level.
96. The method according to claim 81, further comprising:
- h) setting said weighting coefficient of at least one of said markers to approximately zero;
  - i) calculating a value of said objective function with remaining weighting coefficients; and
  - j) determining a contribution of said at least one of said markers to said objective function.
97. The method according to claim 96, further comprising:
- k) repeating steps h) through j) by setting said weighting coefficient of a different at least one of said markers to approximately zero; and
  - l) eliminating a marker from said panel of markers in accordance with said contribution of said marker to said objective function.
98. A method of diagnosing a subject for a disease or condition, comprising:
- a) measuring a value of each of a plurality of markers in said subject;
  - b) calculating a panel response for said subject, said panel response being a function of values of a plurality of markers; and
  - c) diagnosing said disease or condition in said patient when said panel response is greater than a predetermined cutoff.
99. A program product, comprising machine readable program code for causing a machine to perform following method steps:
- a) calculating a value of each of a plurality of markers in a subject;
  - b) calculating a panel response for said subject, said panel response being a function of values of a plurality of markers; and
  - c) diagnosing a disease or condition in said patient when said panel response is greater than a predetermined cutoff.

100. The program product according to claim 99, wherein said machine readable program code is embedded in a portable meter.
101. The program product according to claim 100, wherein said portable meter is a fluorometer.
102. The program product according to claim 100, wherein said portable meter is a reflectometer.
103. The program product according to claim 99, wherein said machine readable program code is embedded in a computer.
104. The program product according to claim 103, wherein said computer is a portable computer.
105. The program product according to claim 103, wherein said computer is adapted to be accessed through a network.
106. The program product according to claim 105, wherein said network is the Internet.
107. The program product according to claim 103, wherein said computer is adapted to be coupled to an analyzer.
108. The program product according to claim 107, wherein said analyzer is an immunoassay analyzer.
109. The program product according to claim 107, wherein said analyzer is a single nucleotide polymorphism detector.
110. The program product according to claim 107, wherein said analyzer is adapted to sort and count similar and different particles and cells.

111. A method of identifying a panel of markers for diagnosis or prognosis of a disease or a condition, comprising:

a) identifying a cutoff region for each of a plurality of markers, said cutoff region being substantially centered about an overlap region of marker values for a set of diseased patients and a set of non-diseased patients, said cutoff region having a location and a length;

b) determining an effectiveness value of each of said plurality of markers in distinguishing said set of diseased patients from said set of non-diseased patients; and

c) defining a panel response as a function of said effectiveness value of each marker and a measured level of each marker.

112. The method according to claim 111, wherein said cutoff region has a length of zero.

113. The method according to claim 111, wherein said cutoff region has a non-zero length.

114. The method according to claim 111, wherein said effectiveness value of each marker is represented by an area under a ROC curve.

115. The method according to claim 3, wherein said indicator function is monotonic with marker value.

116. The method according to claim 115, wherein said indicator function is one of the group consisting of: a ramp function, a step function, and a sigmoid function.

117. The method according to claim 3, wherein said indicator function is adapted to localize a marker value.

118. The method according to claim 117, wherein said indicator function is one of the group consisting of: a triangle, a square, and Gaussian.

119. The method according to claim 1, wherein at least one of said plurality of markers is a derived marker.

120. The method according to claim 119, wherein said derived marker is the ratio of two other markers.

121. The method according to claim 1, wherein said iterating includes using a downhill simplex method.

122. The method according to claim 121, wherein said iterating further includes simulated annealing.

123. The method according to claim 122, wherein said simulated annealing includes performing a statistically sufficient number of optimizations to evaluate a most common solution.

124. The method according to claim 1, wherein said optimization is adapted to provide a stable solution.

125. The method according to claim 124, wherein said adaptation includes varying the marker values by a random percentage.

126. The method according to claim 124, wherein said adaptation includes varying one or more parameters of said panel function.

127. The method according to claim 124, wherein said adaptation includes generating a seed simplex about a minimum.

128. The method according to claim 124, wherein said adaptation includes increasing an annealing temperature until an achieved solution is not recovered.

129. The method according to claim 98, wherein said calculating a panel response includes using a panel response function and parameters relating to said panel response function, said panel response function and parameters being determined by:

- a) calculating a panel response for each patient in a set of diseased patients and in a set of non-diseased patients, said panel response being a function of value of each of a plurality of markers in a panel of markers;
- b) calculating a value for an objective function, said objective function being indicative of an effectiveness of the panel; and
- c) iterating steps a) and b) by varying at least one of parameters relating to said panel response function and a sense of each marker to facilitate optimization of said objective function.

130. The method according to claim 129, wherein said panel of markers is adapted to diagnose two or more diseases or conditions.

131. The method according to claim 130, wherein different panel response functions and parameters are used for each of said diseases or conditions.

132. The method according to claim 98, wherein said at least one of said plurality of markers is a derived marker.

133. The method according to claim 132, wherein said derived marker is a ratio of two other markers.

134. The method according to claim 132, wherein said derived marker is indicative of a change over time in a marker.

135. The method according to claim 98, wherein said calculating a panel response includes calculating, for each patient,  $\sum w_i I_i$ , where  $w$  is a weighting coefficient for a marker  $i$ ,  $I$  is an indicator value for the marker  $i$ , and  $\sum$  is a summation over all of said plurality of markers, said indicator value being a mapping, for each of said plurality of markers, of marker levels, said mapping being according to an indicator function.

136. The method according to claim 135, wherein said indicator function is monotonic with marker value.
137. The method according to claim 136, wherein said indicator function is one of the group consisting of: a ramp function, a step function, and a sigmoid function.
138. The method according to claim 135, wherein said indicator function is adapted to localize a marker value.
139. The method according to claim 138, wherein said indicator function is one of the group consisting of: a triangle, a square, and Gaussian.
140. The method according to claim 20, wherein said removing includes setting a weighting coefficient of said at least one of said markers to approximately zero.
141. The system according to claim 26, wherein said indicator function is monotonic with marker value.
142. The system according to claim 141, wherein said indicator function is one of the group consisting of: a ramp function, a step function, and a sigmoid function.
143. The system according to claim 26, wherein said indicator function is adapted to localize a marker value.
144. The system according to claim 143, wherein said indicator function is one of the group consisting of: a triangle, a square, and Gaussian.
145. The system according to claim 43, wherein said removing includes setting a weighting coefficient of said at least one of said markers to approximately zero.
146. The program product according to claim 49, wherein said indicator function is monotonic with marker value.

147. The program product according to claim 49, wherein said indicator function is one of the group consisting of: a ramp function, a step function, and a sigmoid function.
148. The program product according to claim 49, wherein said indicator function is adapted to localize a marker value.
149. The program product according to claim 49, wherein said indicator function is one of the group consisting of: a triangle, a square, and Gaussian.
150. The program product according to claim 99, wherein said calculating a panel response includes using a panel response function and parameters relating to said panel response function, said panel response function and parameters being determined by:
- a) calculating a panel response for each patient in a set of diseased patients and in a set of non-diseased patients, said panel response being a function of value of each of a plurality of markers in a panel of markers;
  - b) calculating a value for an objective function, said objective function being indicative of an effectiveness of the panel; and
  - c) iterating steps a) and b) by varying at least one of parameters relating to said panel response function and a sense of each marker to facilitate optimization of said objective function.
151. The program product according to claim 150, wherein said panel of markers is adapted to diagnose two or more diseases or conditions.
152. The program product according to claim 151, wherein different panel response functions and parameters are used for each of said diseases or conditions.
153. The program product according to claim 99, wherein said at least one of said plurality of markers is a derived marker.
154. The program product according to claim 153, wherein said derived marker is a ratio of two other markers.

155. The program product according to claim 153, wherein said derived marker is indicative of a change over time in a marker.

156. The program product according to claim 99, wherein said calculating a panel response includes calculating, for each patient,  $\sum w_i I_i$ , where  $w$  is a weighting coefficient for a marker  $i$ ,  $I$  is an indicator value for the marker  $i$ , and  $\sum$  is a summation over all of said plurality of markers, said indicator value being a mapping, for each of said plurality of markers, of marker levels, said mapping being according to an indicator function.

157. The program product according to claim 156, wherein said indicator function is monotonic with marker value.

158. The program product according to claim 157, wherein said indicator function is one of the group consisting of: a ramp function, a step function, and a sigmoid function.

159. The program product according to claim 156, wherein said indicator function is adapted to localize a marker value.

160. The program product according to claim 159, wherein said indicator function is one of the group consisting of: a triangle, a square, and Gaussian.

161. The method according to claim 119, wherein said derived marker is indicative of the change in another marker over time.

162. The method according to claim 119, wherein said derived marker is indicative of the change in said panel response over time.

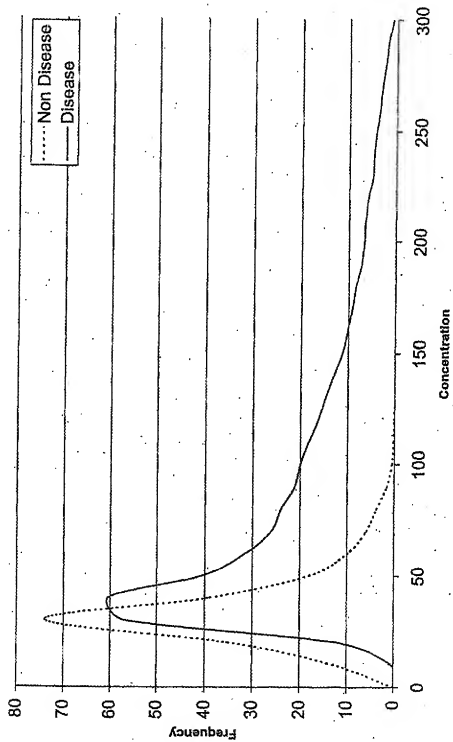
163. The method according to claim 132, wherein said derived marker is indicative of a change over time in said panel response.

164. The program product according to claim 66, wherein said removing includes setting a weighting coefficient of said at least one of said markers to approximately zero.



165. The method according to claim 1, wherein said optimization is adapted to simultaneously at least one of optimize and constrain a plurality of objective functions calculated from a plurality of groups of data.
166. The system according to claim 24, wherein said means for iteratively activating is adapted to simultaneously at least one of optimize and constrain a plurality of objective functions calculated from a plurality of groups of data.
167. The program product according to claim 47, wherein said optimization is adapted to simultaneously at least one of optimize and constrain a plurality of objective functions calculated from a plurality of groups of data.
168. The method according to claim 134, wherein said calculating a panel response is adapted to calculate a panel response without said derived marker when said derived marker is not available and to calculate a panel response using said derived marker when said derived marker is available.
169. The method according to claim 168, wherein said panel response without said derived marker and said panel response using said derived marker are calculated according to different panel response functions and parameters relating to said panel response functions.
170. The method according to claim 163, wherein said calculating a panel response is adapted to calculate a panel response without said derived marker when said derived marker is not available and to calculate a panel response using said derived marker when said derived marker is available.
171. The method according to claim 170, wherein said panel response without said derived marker and said panel response using said derived marker are calculated according to different panel response functions and parameters relating to said panel response functions.

Figure 1 Overlap of Diseased and Non Diseased



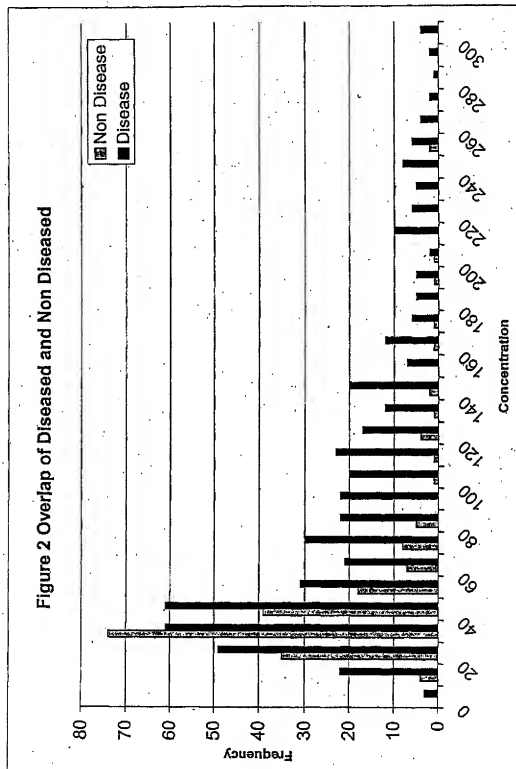


Figure 3: ROC Curve with Area of 0.78

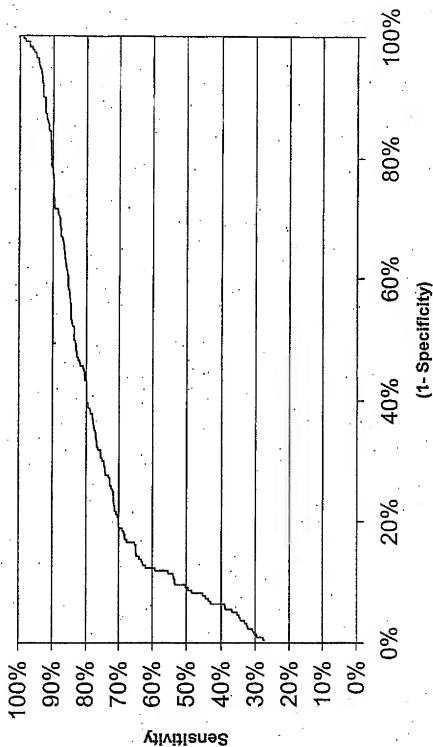
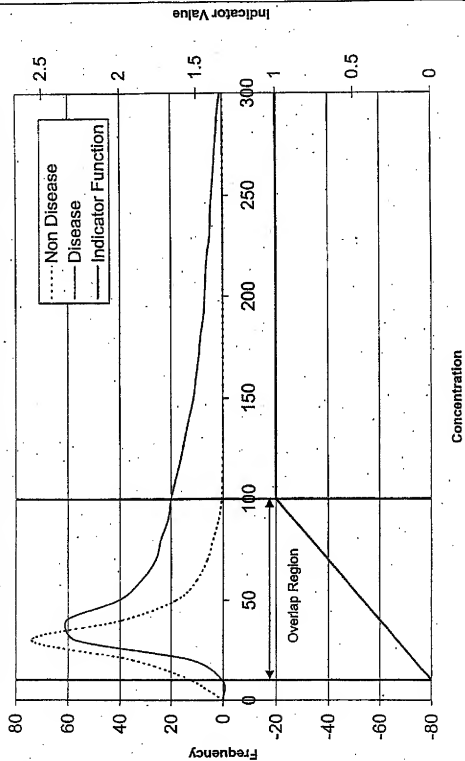


Figure 4 Overlap of Diseased and Non Diseased with Indicator Function



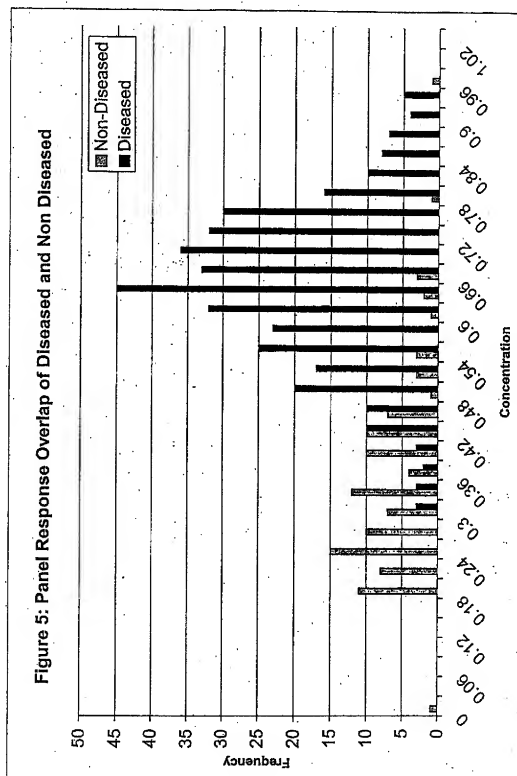


Figure 6: Panel Response ROC Curve with Area of 0.93

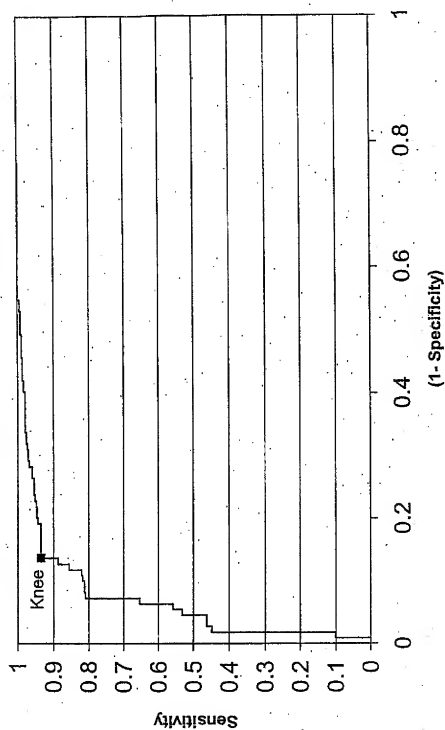
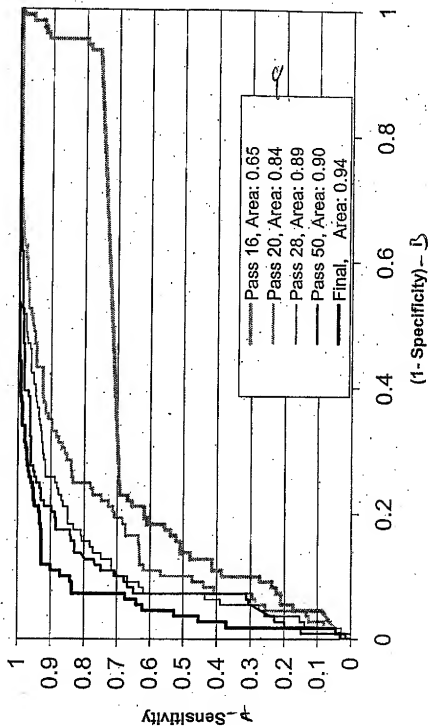


Figure 7: Panel Response ROC Curve Progression During Optimization





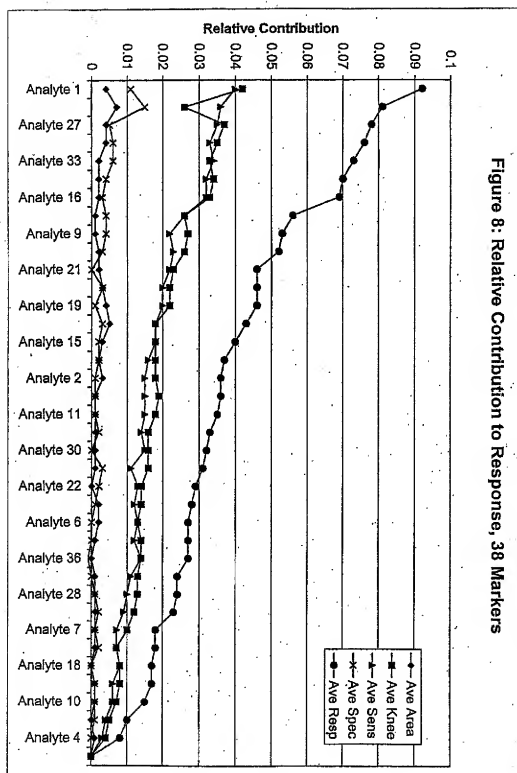


Figure 8: Relative Contribution to Response, 38 Markers

Figure 9: Individual ROC Curves of Five Markers

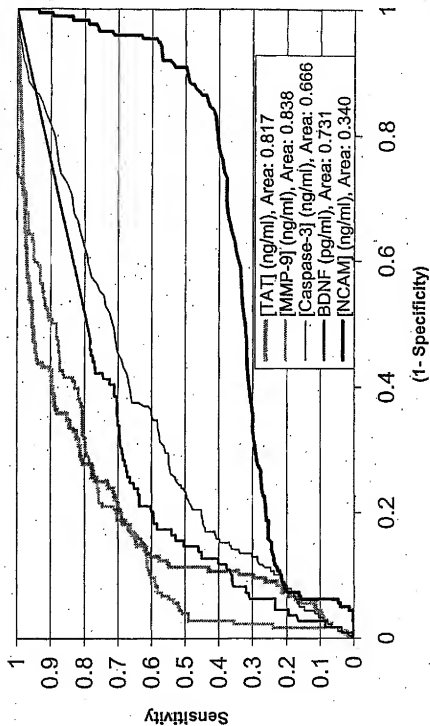


Figure 10: Panel Response ROC Curves, 38 Markers

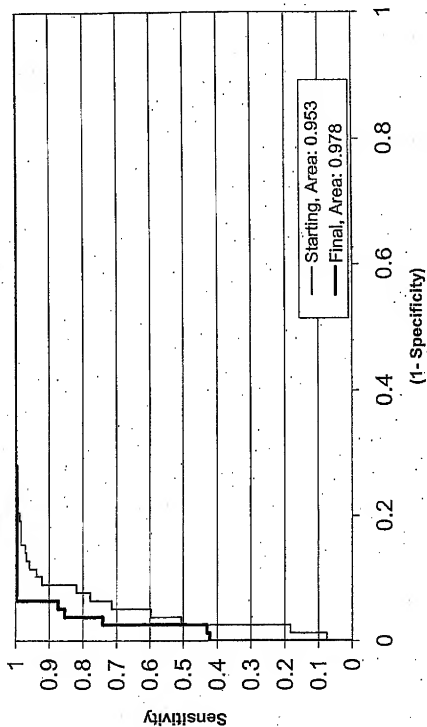


Figure 11: Relative Contribution to Response, 38 Markers

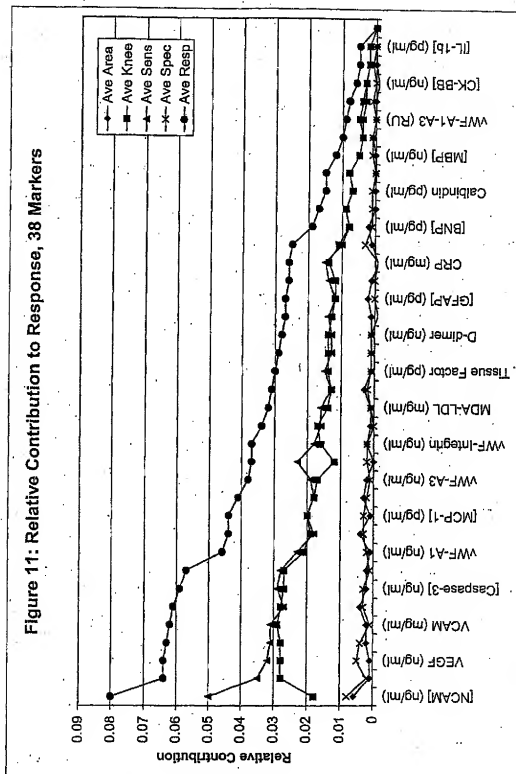


Figure 12: Panel Response ROC Curves, Top 19 Markers

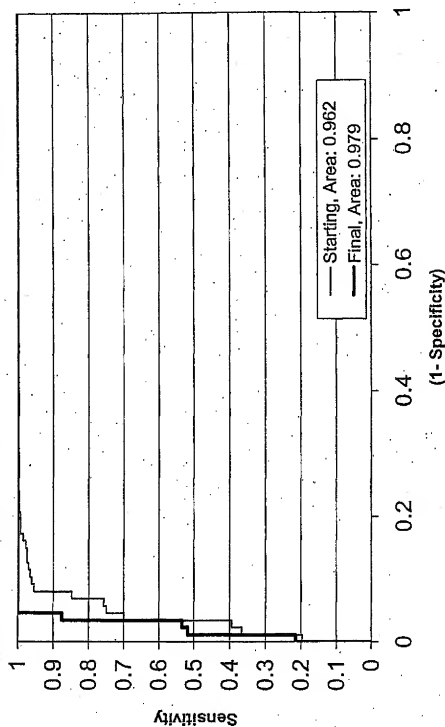


Figure 13: Relative Contribution to Response, Top 19 Markers

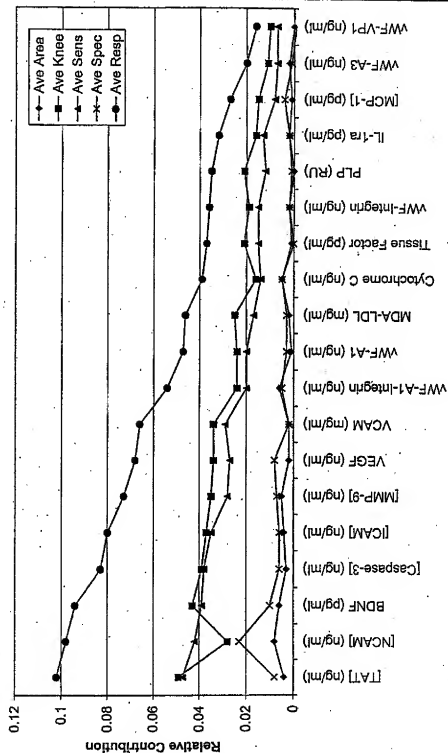


Figure 14: Panel Response ROC Curves, Top 10 Markers

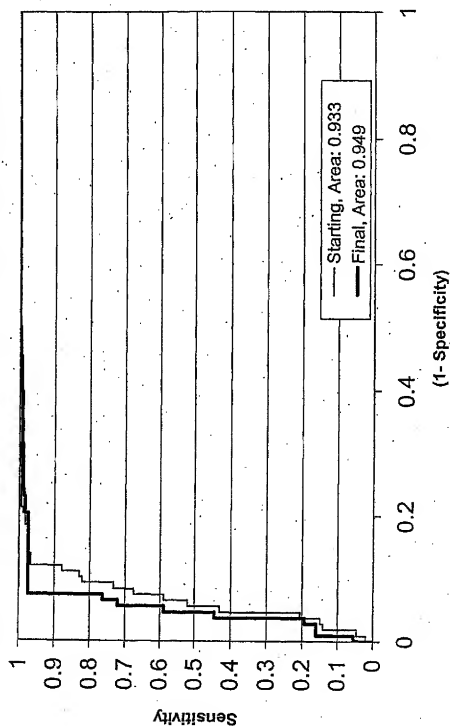


Figure 15: Relative Contribution to Response, Top 10 Markers

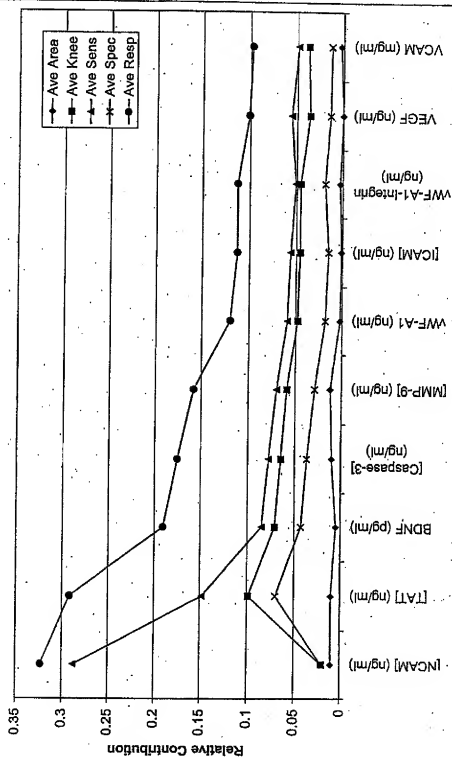




Figure 16: Panel Response ROC Curves, Top 5 Markers

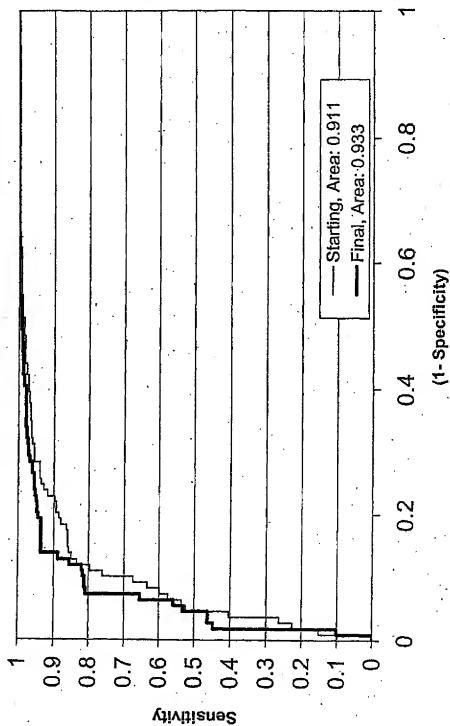


Figure 17: Relative Contribution to Response, Top 5 Markers

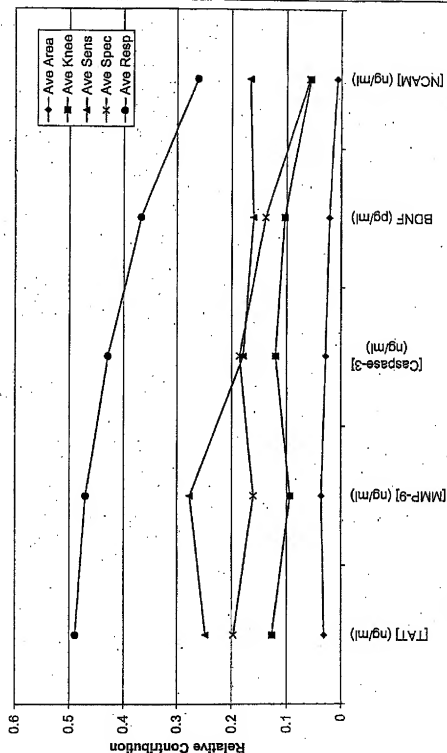


Figure 18: ROC Curves of Panels for AMI

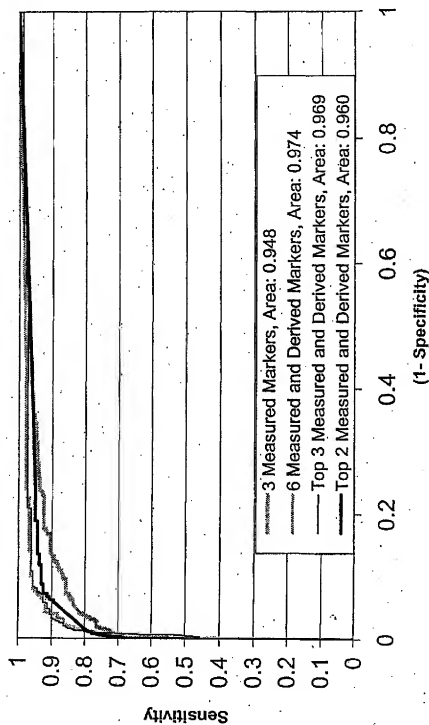


Figure 19: Relative Contribution to Response, 6 AMI Markers

